UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS - ICB
DEPARTAMENTO DE MICROBIOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA
LABORATÓRIO DE VÍRUS

PAULO VICTOR DE MIRANDA BORATTO

A VIROSFERA EM EXPANSÃO: sobre a descoberta, genômica e proteômica de entidades virais codificadoras de um vasto arsenal de proteínas desconhecidas

BELO HORIZONTE
2022

Paulo Victor de Miranda Boratto

A VIROSFERA EM EXPANSÃO: sobre a descoberta, genômica e proteômica de entidades virais codificadoras de um vasto arsenal de proteínas desconhecidas

Tese de Doutorado apresentada ao Programa de Pós-Graduação do Departamento de Microbiologia, Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito final para a obtenção do grau de Doutor em Microbiologia.

Orientador: Prof. Jônatas Santos Abrahão

BELO HORIZONTE

2022

043 Boratto, Paulo Victor de Miranda.
A virosfera em expansão: sobre a descoberta, genômica e proteômica de entidades virais codificadoras de um vasto arsenal de proteínas desconhecidas [manuscrito] / Paulo Victor de Miranda Boratto. – 2022.
103 f. : il. ; 29,5 cm.

Orientador: Prof. Jônatas Santos Abrahão.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. Virologia. 3. Vírus Gigantes. 4. Vírus de DNA. I. Jônatas Santos Abrahão. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 579

Ficha Catalográfica elaborada pela bibliotecária Fabiane C M Reis CRB 6 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
CURSO DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DA DEFESA DE TESE

ATA DA DEFESA DE TESE DE **PAULO VICTOR DE MIRANDA BORATTO**

Nº REGISTRO: 2018698278

Relator e Suplente: Dra. Betânia Paiva Drumond

Às 14:00 horas do dia 02 de fevereiro de 2022, reuniu-se, por via remota, a Comissão Examinadora composta pelos Drs. Jordana Grazziela Alves Coelho dos Reis (Departamento de Microbiologia/ICB /UFMG), Mauricio Lacerda Nogueira (Faculdade de Medicina de São José do Rio Preto), Jaquelline Germano de Oliveira (Instituto René Rachou (IRR)/Fiocruz Minas), Luciana Barros de Arruda (Departamento de Virologia. Instituto de Microbiologia Prof. Paulo de Góes/UFRJ) e o Prof. Jônatas Santos Abrahão - Orientador, para julgar o trabalho final **"A virosfera em expansão: sobre a descoberta, genômica e proteômica de entidades virais codificadoras de um vasto arsenal de proteínas desconhecidos"**, do aluno **Paulo Victor de Miranda Boratto**, requisito final para a obtenção do Grau de **DOUTOR EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Jônatas Santos Abrahão - orientador, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. O candidato foi considerado **APROVADO**. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 02 de fevereiro de 2022.

Membros da banca:

Profa. Dra. Jordana Grazziela Alves Coelho dos Reis
Prof. Dr. Mauricio Lacerda Nogueira
Dra Jaquelline Germano de Oliveira
Profa. Dra. Luciana Barros de Arruda

De acordo:

Prof. Dr. Jônatas Santos Abrahão

(Orientador/ Presidente)

Profa. Dra. Daniele da Glória de Souza

(Coordenadora)

Documento assinado eletronicamente por **Jonatas Santos Abrahao, Professor do Magistério Superior**, em 03/02/2022, às 09:55, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Jordana Grazziela Alves Coelho dos Reis, Professora do Magistério Superior**, em 03/02/2022, às 11:00, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Luciana Barros de Arruda, Usuário Externo**, em 04/02/2022, às 07:07, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Daniele da Gloria de Souza, Coordenador(a) de curso de pós-graduação**, em 10/02/2022, às 08:29, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Jaquelline Germano de Oliveira, Usuário Externo**, em 02/03/2022, às 08:30, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1225224** e o código CRC **1B559200**.

Referência: Processo nº 23072.231004/2021-22 SEI nº 1225224

Agradecimentos:

Neste tópico gostaria de agradecer a todas as instituições e pessoas que me auxiliaram durante essa jornada de quatro anos de doutorado.
Deixo aqui registrado meu agradecimento ao CNPq, CAPES-Cofecub, Fapemig, Ministério da Saúde, Ministério do Meio Ambiente, Secretaria do PPGM, Secretaria da Microbiologia, PRPg-UFMG, ICB-UFMG, PRPq-UFMG e MEC.
Muitas pessoas também me ajudaram, seja de forma direta ou indireta, no desenvolvimento desse trabalho. Gostaria de deixar um agradecimento especial aos meus

pais e família, à Aline, ao meu grande amigo Rayson, ao meu amigo/orientador Jônatas Abrahão e também aos meus amigos de vida Diogo, Erik, Izabela, Karine, Leonardo, Lívia, Lucas, Luís e Pedro. Foram também muito importantes nesse período os colegas e professores do Laboratório de Vírus e também todos os autores dos artigos publicados durante esse período, em especial a Graziele Oliveira, minha parceira de bancada em terras francesas. Por fim, gostaria de agradecer aos professores da Aix-Marseille Université por terem me acolhido, em especial ao professor Bernard La Scola pela ótima orientação, e também aos funcionários das instituições que eu participei.

RESUMO

Recentemente, intensos debates buscaram contribuir a respeito do modo como os vírus são classificados e a maneira como o seu processo evolutivo é interpretado, especialmente após a descoberta dos vírus gigantes de ameba. Esse processo representou uma das fases mais importantes da virologia moderna, tendo seu início em 2003 com a descoberta do Acanthamoeba polyphaga mimivirus (APMV). Posteriormente, impulsionada por aprimoramentos nas técnicas de isolamento viral e cultivo em protistas, uma variedade de vírus de amebas foi encontrada em diferentes ambientes ao redor do globo. Análises de filogenia e reconstruções filogenômicas auxiliaram na classificação desses vírus em um grupo potencialmente monofilético, envolvendo a participação de muitas famílias virais. Esse grupo, denominado de vírus grandes núcleo-citoplasmáticos de DNA (NCLDVs) é atualmente formado por membros das famílias *Poxviridae, Asfarviridae, Iridoviridae, Ascoviridae, Phycodnaviridae, Marseilleviridae* e *Mimiviridae*, além das linhagens representadas pelos pithovírus, pandoravírus, mollivírus, medusavírus, pacmanvírus, faustovírus, klosneuvírus, entre outros. Mesmo potencialmente tendo compartilhado um ancestral em comum, membros do NCLDV apresentam importantes diferenças morfogenéticas que estão em plena expansão, visto a quantidade crescente de entidades virais isoladas e/ou detectadas por metodologias independentes de cultivo. Neste trabalho, o nosso objetivo foi caracterizar amostras de vírus de amebas *Acanthamoeba castellanii* obtidas a partir de diferentes ambientes brasileiros. Para isso, foram feitas caracterizações genômicas, de proteômica e/ou biológicas de três linhagens evolutivas de vírus (tupanvírus, pandoravírus e yaravírus) isoladas desse hospedeiro amebiano. A análise de proteômica das partículas dos tupanvírus revelou a presença de mais de 100 proteínas codificadas por genes virais, a maioria sem função ou origem conhecida. Neste contexto, cerca de 20% das proteínas encontradas são codificadas por genes que foram transferidos de organismos celulares para os tupanvírus ao longo de sua história evolutiva. Desses 20%, cerca de 9% são provenientes de eucariotos (um terço originário de amebas), 3% provenientes de arqueias e os outros 8% provenientes de bactérias. Com relação aos outros 80% do proteoma do tupanvírus soda-lake, a maior parte encontra-se relacionada especialmente com outros membros do NCLDV.

Para os pandoravírus, a presente tese contribui para um melhor entendimento do seu ciclo de multiplicação. Nossos dados indicam que todos os isolados analisados são capazes de modificar intensamente o ambiente citoplasmático das amebas infectadas, recrutando mitocôndrias e membranas que são aproveitadas na formação das fábricas virais. Também foram identificados diferentes padrões de morfogênese das partículas, com a montagem podendo ser iniciada tanto pelo ápice como pela base do vírion. Por meio da contagem de partículas durante o ciclo, observou-se que os pandoravírus podem sofrer exocitose logo após a morfogênese, em um processo que envolve um intenso recrutamento de membranas celulares para as fábricas virais. Tratando células infectadas com brefeldina, a exocitose das partículas virais foi afetada em duas das três linhagens analisadas, o que pode indicar uma variabilidade entre os isolados. Apesar da exocitose, a lise de células hospedeiras também se mostrou um importante fator para a liberação viral. Por fim, com relação ao yaravírus, nossos dados reportam a descoberta de uma linhagem viral inédita, apresentando uma origem e filogenia intrigantes. Com partículas de aproximadamente 80 nm e um genoma de cerca de 45 kpb, o yaravírus passou a representar o menor e menos complexo (em termo de quantidade de genes) vírus isolado de *Acanthamoeba spp*. Mais de 90% do genoma do yaravírus não apresenta similaridade a sequências de bancos de dados de domínios públicos. Nossos dados também descrevem a identificação de uma proteína principal de capsídeo com sequência de aminoácidos divergente, sem nenhuma homologia significativa com proteínas de capsídeo observadas para outros membros do grupo NCLDV. No entanto, esta proteína apresenta.uma estrutura predita para o domínio de *double-jelly roll*. Além disso, por meio de buscas por similaridade de sequências em bancos de dados públicos contendo mais de 8.500 metagenomas, foi encontrada apenas homologias muito distantes para a proteína ATPase viral

(~33% de similaridade), o que destaca a raridade desse vírus no ambiente natural e demonstra quão importantes se fazem os estudos baseados em isolamento, e não apenas os voltados para detecção genômica. Concluindo, nosso trabalho não só permitiu expandir o atual conhecimento sobre a diversidade de vírus de eucariotos como também proveu informações que podem desafiar a atual classificação dos vírus de DNA.

Palavras-chave: vírus gigantes, tupanvírus, pandoravírus, yaravírus

ABSTRACT

Recently, intense debates have sought to contribute in the way viruses are classified and the way their evolutionary process is interpreted, especially after the discovery of the giant viruses of amoeba. This process represented one of the most important phases of modern virology, beginning in 2003 with the discovery of Acanthamoeba polyphaga mimivirus (APMV). After, driven by major improvements in the techniques of viral isolation and cultivation in protists, a variety of amoeba viruses were found in different environments around the globe. Phylogeny analyzes and phylogenomic reconstructions ended up classifying these viruses into a potentially monophyletic group, involving the participation of many viral families. This group, called nucleocytoplasmic large DNA viruses (NCLDVs), is currently formed by members of the families *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Marseilleviridae* and *Mimiviridae*, in addition to the strains represented by the pithovirus, pandoravirus, mollivirus, medusavirus, pacmanvirus, faustovirus, klosneuvirus, among others. Although they potentially shared a common ancestor, members of the NCLDV have important morphogenetic differences that are in full expansion, given the increasing number of isolated and/or detected viral entities by recently incorporated cultivation-independent methods. In this work, our aim is to characterize samples of viruses from the amoeba *Acanthamoeba castellanii*, obtained from different Brazilian environments. For that, we promoted the genomic, proteomic and/or biological characterizations of three evolutive viral strains (tupanvirus, pandoravirus and yaravirus) isolated from this amoebal host. Proteomic analysis of tupanvirus particles revealed the presence of more than 100 proteins encoded by viral genes, a great part of them without known function or origin. In this context, about 20% of the proteins are encoded by genes that were transferred from cellular organisms to tupanviruses throughout their evolutionary history. For pandoraviruses, the present thesis has contributed to a better understanding of the multiplication cycle of the members of this lineage. Our data indicates that all the analyzed isolates are capable of deeply modify the cytoplasmic environment of the infected amoebae, recruiting mitochondria and membranes that are used in the formation of viral factories. Various patterns of pandoravirus particle morphogenesis were observed, and the assembly seemed to be started by the apex or the base of the particle. By counting the viral particles during the infection cycle, we observed that pandoraviruses can undergo exocytosis shortly after morphogenesis, in a process that involves intense recruitment of membranes surrounding the newly formed particles. Treating the infected cells with brefeldin, the exocytosis of viral particles was affected in two of the three analyzed strains, indicating variability between the isolates. Despite occurrence of exocytosis, lysis of host cells proved to be an important contributor to the viral release. Finally, regarding the analyzes made for the yaravirus, our data report the discovery of a viral lineage never described before, presenting a puzzling origin and phylogeny. With particles of approximately 80nm and a genome of about 45kbp, the yaravirus now represents the smallest and simplest (in terms of number of genes) virus isolated from *Acanthamoeba spp*. Interestingly, more than 90% of its genome does not show sequence similarity in databases. Our data also describe the

identification of a divergent major capsid protein, with no significant homology to capsid proteins seen in other NCLDV members, but with a predicted structure for the double-jelly roll domain. In addition, only distant-related ATPase homologs were found in similarity searches for sequences deposited in more than 8,500 public metagenomes, highlighting the rarity of this virus in the natural environment and demonstrating how important are studies based on viral isolation, not just those focused on genome detection. In conclusion, our work not only allows us to expand the current knowledge about the diversity of eukaryotic viruses, but also complements it with information that may challenge the current classification of the DNA viruses.

Keywords: giant virus, tupanvirus, pandoravirus, yaravirus

SUMÁRIO

1. INTRODUÇÃO

Quase duas décadas já se passaram desde a descoberta e caracterização do primeiro vírus gigante de ameba. Naquele momento, um passo importante foi dado em direção a um maior entendimento do ambiente microbiológico ao qual estamos inseridos, auxiliando também em uma grande expansão do nosso conhecimento sobre a virosfera. Os vírus a partir de então pararam de ser vistos exclusivamente como entidades muito simples e passaram a ser aceitos como organismos capazes de apresentar genomas de alta complexidade e tamanhos de partícula inesperadamente grandes, podendo estar associados nos mais diversos tipos de relações evolutivas que datam desde o surgimento da vida na Terra. Posteriormente, impulsionada por aprimoramentos nas técnicas de isolamento viral e cultivo em protistas, uma grande variedade desses vírus foi encontrada em diferentes ambientes ao redor do globo. Análises de filogenia e reconstruções filogenômicas acabaram por auxiliar na classificação dessas entidades em um grupo potencialmente monofilético, envolvendo a participação de muitas famílias virais. Esse grupo, denominado de vírus grandes núcleo-citoplasmáticos de DNA (NCLDVs), é atualmente formado por membros das famílias *Poxviridae, Asfarviridae, Iridoviridae, Ascoviridae, Phycodnaviridae, Marseilleviridae* e *Mimiviridae*, além das linhagens representadas pelos pithovírus, pandoravirus, mollivírus, medusavírus, pacmanvírus, faustovírus, klosneuvírus, entre outros. Mesmo compartilhando entre si um potencial ancestral em comum, os NCLDV apresentam importantes diferenças morfogenéticas, cada vez mais evidentes devido à crescente quantidade de novos grupos virais isolados e/ou detectados por metodologias independentes de cultivo.

Durante os últimos dez anos o nosso grupo de pesquisa tem trabalhado para contribuir para o conhecimento sobre a diversidade e excepcionalidade dos vírus gigantes de ameba, utilizando para isso de amostras ambientais coletadas no Brasil, um país com uma das maiores biodiversidades do planeta. Dezenas de trabalhos científicos já foram gerados durante esse período, apresentando como um de seus principais focos o isolamento e a caracterização de amostras de vírus gigantes brasileiros obtidas em várias regiões do país. Um fator de extrema importância para esse sucesso foi também a parceria consolidada durante esse tempo com pesquisadores da Universidade de Aix-Marseille, na França, em especial os Prof. Bernard La Scola, sendo ele uma importante referência no assunto e pioneiro no mundo na descrição de entidades virais tão complexas. Com o passar dos anos, várias parcerias com instituições nacionais e internacionais também foram estabelecidas, incluindo a Universidade Federal do Rio de Janeiro, a Universidade de Brasilia, a EMBRAPA Pantanal, Petrobrás, Purdue

University (EUA) e o Lawrence Berkeley National Laboratory (EUA), os quais foram essenciais para o desenvolvimento dos trabalhos do grupo. A presente tese de doutorado foi construída então sobre fortes bases estabelecidas pelo nosso grupo durante a última década, apresentando como apelo a expansão do nosso conhecimento sobre os vírus gigantes em uma série de estudos que buscaram atingir diversas vertentes.

Por meio de análises de proteômica e a construção de árvores filogenéticas, pudemos investigar melhor sobre as origens das 127 proteínas que formam as partículas dos tupanvírus; esse que atualmente é um dos vírus de maior tamanho e complexidade genética já observados na natureza. Em um segundo estudo, por meio de uma metodologia voltada especialmente para a utilização de ensaios de microscopia eletrônica de transmissão e varredura (MET e MEV, respectivamente), conseguimos reunir um acervo de mais de 200 imagens que nos permitiu estabelecer uma descrição detalhada do ciclo de multiplicação de três amostras de pandoravírus isoladas em diferentes locais do Brasil. Essa descrição envolveu não só os estágios básicos do processo de infecção viral, mas também o modo como o hospedeiro celular se comporta durante cada etapa do ciclo, além de algumas diferenças biológicas vistas durante o processo de infecção dessas amostras. Por fim, finalizamos esse trabalho com a descrição do isolamento e caracterização do yaravírus, uma nova entidade viral que apresenta uma das origens evolutivas e filogenias mais intrigantes já observadas em um vírus de ameba. Uma abordagem descritiva bastante completa foi realizada nesse sentido. Foram desenvolvidos conjuntos de ensaios biológicos e genéticos organizados em diversas vertentes, como, por exemplo: (i) análises do vírus e de seu ciclo de infecção por MET, MEV, microscopia crio-eletrônica, tomografia eletrônica e coloração negativa; (ii) análises da replicação do material genético viral por PCR quantitativa; (iii) caracterizações genéticas e morfológicas das partículas purificadas do yaravírus por meio de experimentos de sequenciamento do genoma completo, predição e anotação gênica, além de análises de proteômica da partícula madura e, por fim, (iv) o estabelecimento de importantes relações filogenéticas dessa linhagem viral em meio a um conjunto de outros vírus de ameba com características muito diferentes. Todo esse processo gerou como consequência a criação de uma nova família para acomodar outros vírus da linhagem do yaravírus que vierem a ser descobertos. Portanto, o presente trabalho de tese lidou com importantes desafios, uma vez que as entidades virais aqui estudadas codificam para um vasto arsenal de proteínas completamente desconhecidas, não apresentando todos os genes de referência utilizados para a megataxonomia recentemente proposta pelo Comitê Internacional de Taxonomia dos Vírus (ICTV). Como conclusão, essa tese de doutorado não só permitiu expandir o atual

conhecimento sobre a diversidade de vírus de eucariotos, em especial os vírus de protistas, como também contribuiu para gerar interessantes discussões sobre a megataxonomia viral.

Os estudos desenvolvidos durante a presente tese foram executados, parcial ou integralmente, de acordo com os seguintes cadastros:

- SISGEN - AA3B21E, A702EB8, A25764F, AC31840, A473BD3, A3DAB3F, AC3045D, A96431C, ABF23CC, A2F8816, A580BBD e AEC3EAA (acesso ao patrimônio genético brasileiro);
- SISBIO - 33326, 34293 e 80252 (coletas de material microbiológico para prospecções).

2. Artigo número 01:

2.1 The multiple origins of proteins present in tupanvirus particles

(doi: 10.1016/j.coviro.2019.02.007)

Nesse artigo fizemos inicialmente a compilação de todos os dados de proteômica disponíveis para os vírus gigantes de ameba estudados até o período de junho de 2019. Foram incluídas informações de trabalhos envolvendo os vírus Acanthamoeba polyphaga mimivirus, Cafeteria roenbergensis virus, diferentes isolados de pandoravírus, faustovírus E12, pithovírus sibericum, mollivírus sibericum e tupanvírus soda-lake.

Baseando-se em um contexto daquele momento, os vírus conhecidos como tupanvírus haviam sido recentemente descobertos pelo nosso grupo de pesquisa, tendo ganhado um notável destaque internacional por sua complexidade genética e morfologia peculiar. Da mesma forma, a proteômica das partículas maduras revelou que os tupanvírus são compostos por um conjunto de centenas de proteínas, muitas delas com função completamente desconhecida e sem homologia com sequências de proteínas disponíveis em bancos de dados (condificadas por genes conhecidos como ORFans). Tendo isso em mente, neste artigo buscamos gerar e analisar as relações filogenéticas de proteínas que sabidamente são responsáveis por formar as partículas do tupanvírus soda-lake, comparando com sequências de proteínas similares de membros de diferentes grupos taxonômicos. Para isso, para cada uma das 127 proteínas que formam as partículas purificadas do tupanvírus foram também analisadas as similaridades de suas sequências com as dos membros do grupo Eukarya (30 *best-hits*), Archaea (30 *best-hits*), Amoebozoa (10 *best-hits*), outros vírus (10 *best-hits*), Proteobacteria (15 *best-hits*) e Firmicutes (15 *best-hits*). As sequências de proteínas foram então coletadas, alinhadas e as árvores filogenéticas foram construídas com base no método de máxima verossimilhança, em uma análise de *bootstrap* contendo 1000 replicatas. As árvores geradas demonstraram que cerca de 20% do proteoma das partículas dos tupanvírus se agrupam com membros do grupo dos eucariotos, arqueias e bactérias. Desses 20%, cerca de 9% são provenientes de eucariotos (um terço originário de amebas), 3% provenientes de arqueias e os outros 8% provenientes de bactérias. Com relação aos outros 80% do proteoma do tupanvírus soda-lake, a maior parte pode se encontrar relacionada especificamente (mas não exclusivamente) com outros membros do NCLDV. Por fim, como uma grande porção de genes foi observada sendo compartilhada com membros de outros domínios celulares, nossos dados sugerem que os tupanvírus adquiriram genes de diferentes organismos ao longo de sua evolução e incorporaram algumas das proteínas codificadas por estes genes em suas partículas virais.

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 Cortines** and **Peter Prevelige**

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].

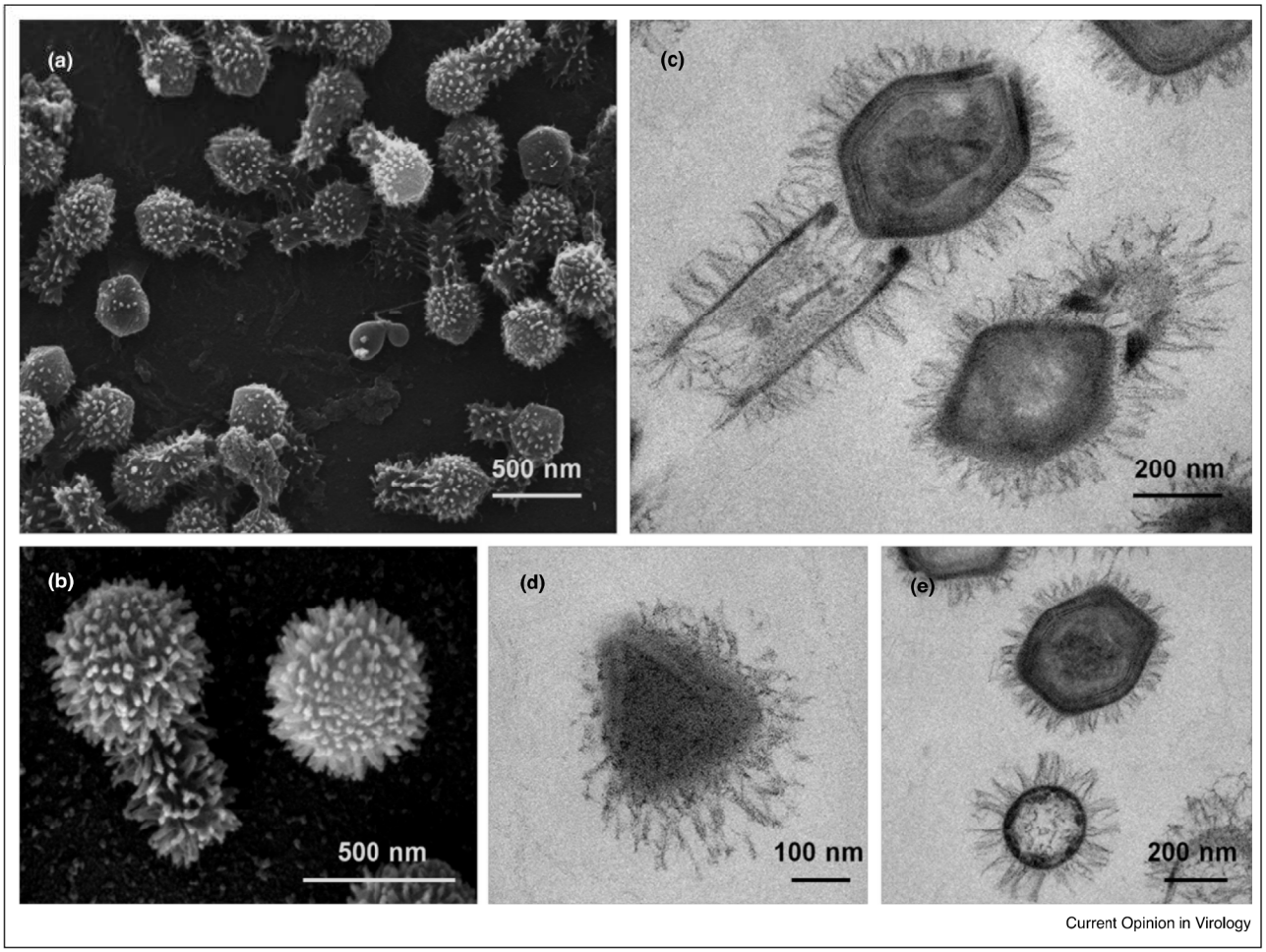
Even more striking is the structure observed for tupanviruses, a new group of viruses within the family *Mimiviridae*. These viruses were recently isolated from extreme environments in Brazil and are capable of infecting a wide variety of amoebae species [22••]. Tupanvirus has more than 1200 genes, and 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

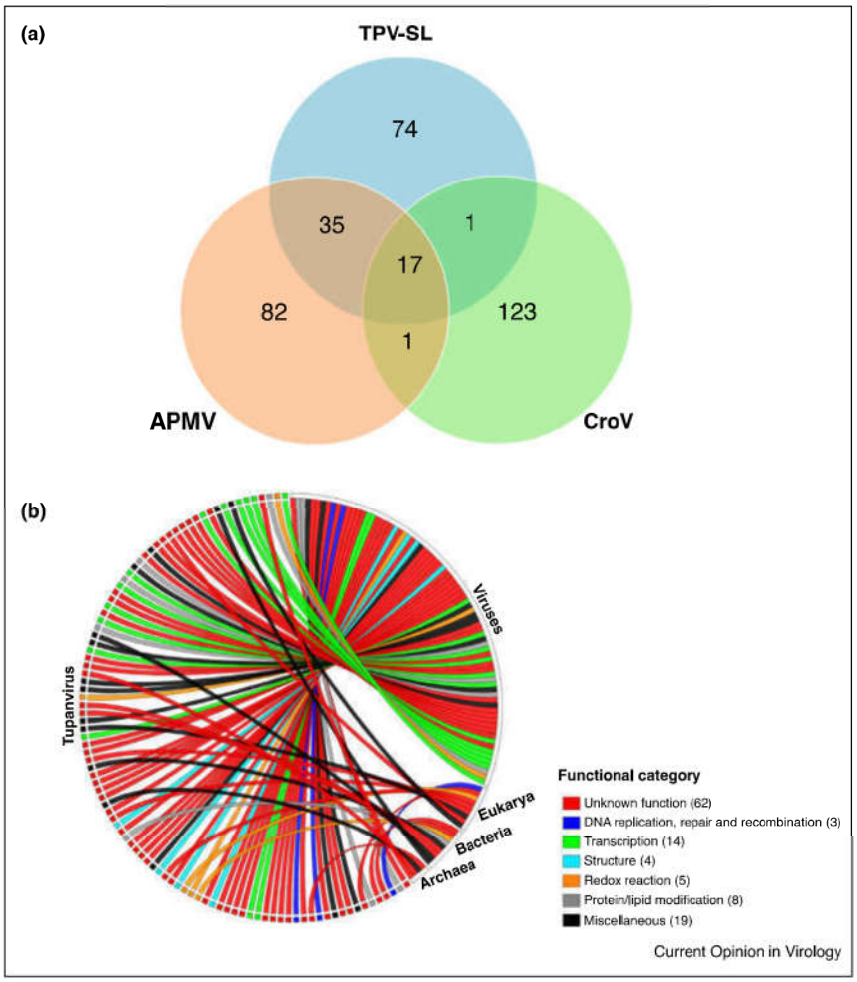
According to many proteomic studies within the known virosphere, the mature particles of amoebal giant viruses are composed of a great number of proteins [4,24••,25,26•]. This has generated hypotheses that seek to explain why only a small fraction of the nearly one thousand proteins encoded by mimiviruses 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



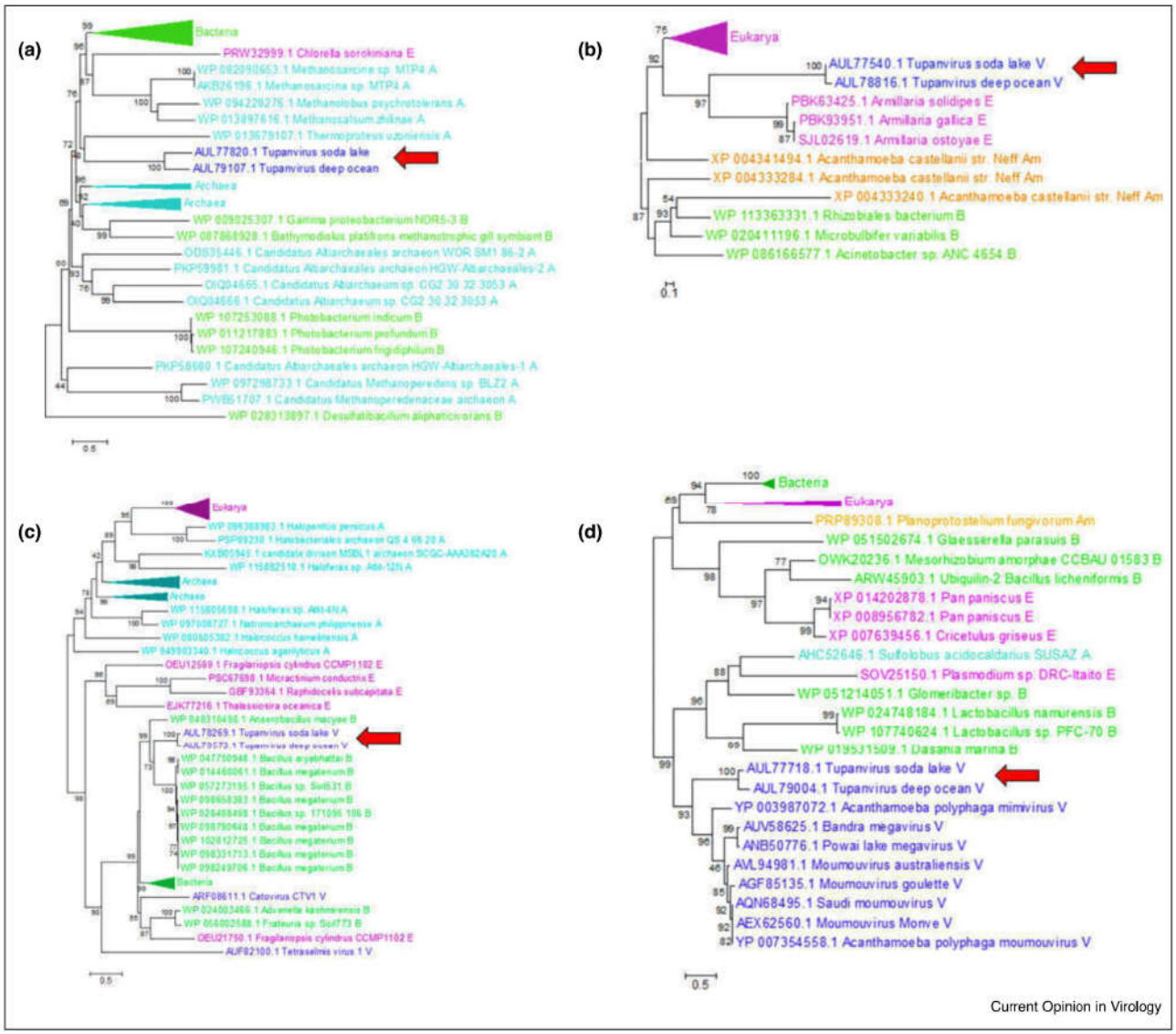
Proteome analysis of tupanvirus soda lake.
(a) Venn diagram of a comparative analysis of the viral particle proteome of TPV-SL, APMV, and CroV containing 127, 136, and 141 proteins, respectively. The analysis was performed using the proteomic data of each virus obtained from the literature [21,25,31] 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.

3. Artigo número 02:

3.1 New isolates of pandoraviruses: contribution to the study of replication cycle steps

(doi: 10.1128/JVI.01942-18)

A potencial família *Pandoraviridae* é composta por um dos vírus mais complexos já observados na natureza até o momento. Apesar disso, são poucos os isolados descritos atualmente, o que faz com que muitos aspectos do seu ciclo precisem ser mais bem elucidados. O artigo a seguir foi desenvolvido nesse contexto, aproveitando-se de uma série de metodologias de imagem com o objetivo de gerar uma descrição detalhada do ciclo de multiplicação de três amostras de pandoravírus brasileiras, o pandoravírus kadiweu, o pandoravírus pampulha e o pandoravírus tropicalis. O primeiro isolado foi obtido por meio de amostras de água coletadas na cidade de Bonito, Mato Grosso do Sul, seguida da aplicação de metodologias de cultivo em amebas já bem estabelecidas na literatura. As outras duas amostras faziam parte de uma coleção de isolados de vírus gigantes obtidas em um estudo de prospecção realizado a partir de de 2015, tendo sido coletadas a partir de amostras de esgoto da lagoa da Pampulha, Belo Horizonte. Após o isolamento, cada uma das amostras foi confirmada como um isolado de pandoravírus po meio de observação da partícula por MET e sequenciamento de um fragmento do gene conservado da DNA polimerase viral. Posteriormente, cada uma dessas amostras foi utilizada em um extenso experimento de infecção assincrônica em amebas da espécie *Acanthamoeba castellanii*, culminando em uma coleção de mais de 200 imagens de microscopia eletrônica de transmissão e de varredura. Essas imagens foram intensamente analisadas e utilizadas para a caracterização de diversas etapas do ciclo de multiplicação dessas amostras virais. Observou-se que, assim como já descrito para outros vírus gigantes, todas as amostras de pandoravírus parecem penetrar na célula hospedeira por meio de eventos de fagocitose. Esses vírus são então internalizados dentro de um fagossomo até que este se funda com organelas lisossomais e esse processo desencadeie o desnudamento da partícula, com liberação do genoma no citoplasma da célula hospedeira. Em uma etapa posterior, grandes regiões elétron-lucentes parecem ser produzidas na forma de fábricas virais, contendo um grande número de mitocôndrias e outras organelas, potencialmente auxiliando na morfogênese das partículas. A dinâmica de morfogênese então se inicia, com as partículas sendo formadas inicialmente por meio de crescentes e adquirindo com o tempo tanto um maior tamanho quanto complexidade. Por fim, antes de ocorrer o processo de lise celular, como previamente descrito em outros estudos, as partículas de pandoravírus parecem ser liberadas da célula hospedeira por exocitose, apesar de que diferenças biológicas vistas entre as amostras. Como conclusão,

esse trabalho contribuiu de maneira importante para um melhor entendimento do ciclo de multiplicação dos pandoravírus.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

New Isolates of Pandoraviruses: Contribution to the Study of Replication Cycle Steps

Ana Cláudia dos Santos Pereira Andrade,[a] Paulo Victor de Miranda Boratto,[a] Rodrigo Araújo Lima Rodrigues,[a] Talita Machado Bastos,[a] Bruna Luiza Azevedo,[a] Fábio Pio Dornas,[b] Danilo Bretas Oliveira,[b] Betânia Paiva Drumond,[a] Erna Geessien Kroon,[a] Jônatas Santos Abrahão[a]

[a]Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[b]Universidade Federal do dos Vales do Jequitinhonha e Mucuri, Diamantina, Brazil

ABSTRACT Giant viruses are complex members of the virosphere, exhibiting outstanding structural and genomic features. Among these viruses, the pandoraviruses are some of the most intriguing members, exhibiting giant particles and genomes presenting at up to 2.5 Mb, with many genes having no known function. In this work, we analyzed, by virological and microscopic methods, the replication cycle steps of three new pandoravirus isolates from samples collected in different regions of Brazil. Our data indicate that all analyzed pandoravirus isolates can deeply modify the *Acanthamoeba* cytoplasmic environment, recruiting mitochondria and membranes into and around the electron-lucent viral factories. We also observed that the viral factories start forming before the complete degradation of the cellular nucleus. Various patterns of pandoravirus particle morphogenesis were observed, and the assembly of the particles seemed to be started either by the apex or by the opposite side. On the basis of the counting of viral particles during the infection time course, we observed that pandoravirus particles could undergo exocytosis after their morphogenesis in a process that involved intense recruitment of membranes that wrapped the just-formed particles. The treatment of infected cells with brefeldin affected particle exocytosis in two of the three analyzed strains, indicating biological variability among isolates. Despite such particle exocytosis, the lysis of host cells also contributed to viral release. This work reinforces knowledge of and reveals important steps in the replication cycle of pandoraviruses.

IMPORTANCE The emerging Pandoraviridae family is composed of some of the most complex viruses known to date. Only a few pandoravirus isolates have been described until now, and many aspects of their life cycle remain to be elucidated. A comprehensive description of the replication cycle is pivotal to a better understanding of the biology of the virus. For this report, we describe new pandoraviruses and used different methods to better characterize the steps of the replication cycle of this new group of viruses. Our results provide new information about the diversity and biology of these giant viruses.

KEYWORDS pandoravirus, giant virus, replication cycle, viral morphogenesis, viral release, virus diversity

Giant viruses are a group of complex viruses commonly referred to as nucleocytoplasmic large DNA viruses (NCLDV); the members of the group exhibit diverse characteristics that have been astonishing the scientific community over the last few years. Different groups of viruses described to date are able to replicate in amoeba cells, expanding considerably our knowledge about their diversity, structure, genomics, and evolution (1–5).

Five years ago, two complex giant viruses infecting *Acanthamoeba castellanii* cells

Citation Pereira Andrade ACDS, Victor de Miranda Boratto P, Rodrigues RAL, Bastos TM, Azevedo BL, Dornas FP, Oliveira DB, Drumond BP, Kroon EG, Abrahão JS. 2019. New isolates of pandoraviruses: contribution to the study of replication cycle steps. J Virol 93:e01942-18. https://doi.org/10.1128/JVI.01942-18.

Editor Rozanne M. Sandri-Goldin, University of California, Irvine

Address correspondence to Jônatas Santos Abrahão, jonatas.abrahao@gmail.com.

A.C.D.S.P.A. and P.V.D.M.B. contributed equally to this article.

Received 1 November 2018
Accepted 1 November 2018
Accepted manuscript posted online 12 December 2018
Published 19 February 2019

were described, constituting a new group of viruses called pandoraviruses. One of the isolated viruses, which originated from a marine sediment layer of the Tunquen River in Chile, was named *Pandoravirus salinus*, and the other one, isolated from the mud of a freshwater pond in Australia, was named *P. dulcis*. Pandoraviruses have morphological and genetic characteristics that have never been described before, such as an oval-shaped particle with an ostiole-like apex, measuring ~1.0 μm in length and ~0.5 μm in diameter, representing some of the largest viruses known to date (6). These viruses are also marked by the presence of a double-stranded DNA genome of up to 2.5 Mb (*P. salinus*), currently the largest genome in the virosphere (6).

In 2008, amoebas of the *Acanthamoeba* genus harboring an unknown endocytobiont were isolated from the contact lens and inflamed eye of a patient with keratitis in Germany (7). Years after this discovery, analysis of this endosymbiont genome revealed the viral nature of this organism, which was classified as a pandoravirus (8). This was the third pandoravirus described, and it was named *P. inopinatum* (9, 10). In 2015 to 2016, new pandoraviruses were described using a culture of *A. castellanii* cells belonging to sewage and soda lake water samples. These viruses were named *P. massiliensis*, *P. pampulha*, and *P. brasiliensis* (11–13). Another recent prospective study reported the isolation of *Pandoravirus quercus*, isolated from samples of soil collected in Marseille (France); *P. neocaledonia*, isolated from the brackish water around a mangrove near Noumea Airport (New Caledonia), and *P. macleodensis*, isolated from a freshwater pond near Melbourne (Australia) (14). Pandoraviruses represent a genome exceeding those of some eukaryotic microorganisms, with a huge proportion of open reading frame (ORF) genes without homologs (ORFans) in any database. The ORFans correspond to about 70% of the predicted genes of pandoraviruses (6).

Despite the plethora of novel characteristics revealed by analyses of the genomes and evolution of the pandoraviruses, their replication cycle still needs further study to be better understood. In the present report, we present an in-depth investigation of the replication cycle steps of three new isolates of pandoraviruses. We observed that the pandoraviruses are able to deeply modify the acanthamoeba cytoplasmic environment, recruiting mitochondria and membranes into and around the electron-lucent viral factories (VFs). The viral factory formation and viral particle morphogenesis were analyzed in an in-depth manner by electron microscopy (EM), with results reinforcing previously published data and revealing new features about pandoraviruses' replication cycles. We also demonstrated by microscopy and pharmacological inhibition of membrane traffic that viral particles were released from infected cells both by exocytosis and by cell lysis. This work contributes to the understanding of important steps in the replication cycle of pandoraviruses.

RESULTS

New members of the emerging family Pandoraviridae. Isolation of a new pandoravirus isolate, namely, *P. kadiweu*, was performed by culturing amoebas of the *A. castellanii* species with water samples collected in the city of Bonito, Mato Grosso do Sul, Brazil (Fig. 1A, D, and H). A prospective study conducted between 2015 and 2017 using culture of *A. castellanii* species with sewage samples from different environmental and clinical samples reported the collection of two pandoravirus isolates that were identified by real-time PCR and electron microscopy (12). The pandoravirus isolates were obtained from samples of Mergulhão Creek and Bom Jesus Creek, in the region of Pampulha Lake, Belo Horizonte, Brazil (Fig. 1A to C), and were named *Pandoravirus pampulha* (12) (Fig. 1F) and *P. tropicalis*, respectively (Fig. 1E and G). The *P. kadiweu* isolate and the two isolates described by Andrade et al. in 2018 (12) are new members of the emerging family Pandoraviridae.

The isolates were observed both by optical microscopy (data not shown) and by electron microscopy, and the images indicated no evident morphological differences among the three isolates (Fig. 1E to H). The isolates were ~1.0 μm in length and had an ostiole-like apex at one end of the particle as previously described for other pandoraviruses (6, 12–15). In order to evaluate whether our isolates were similar, we

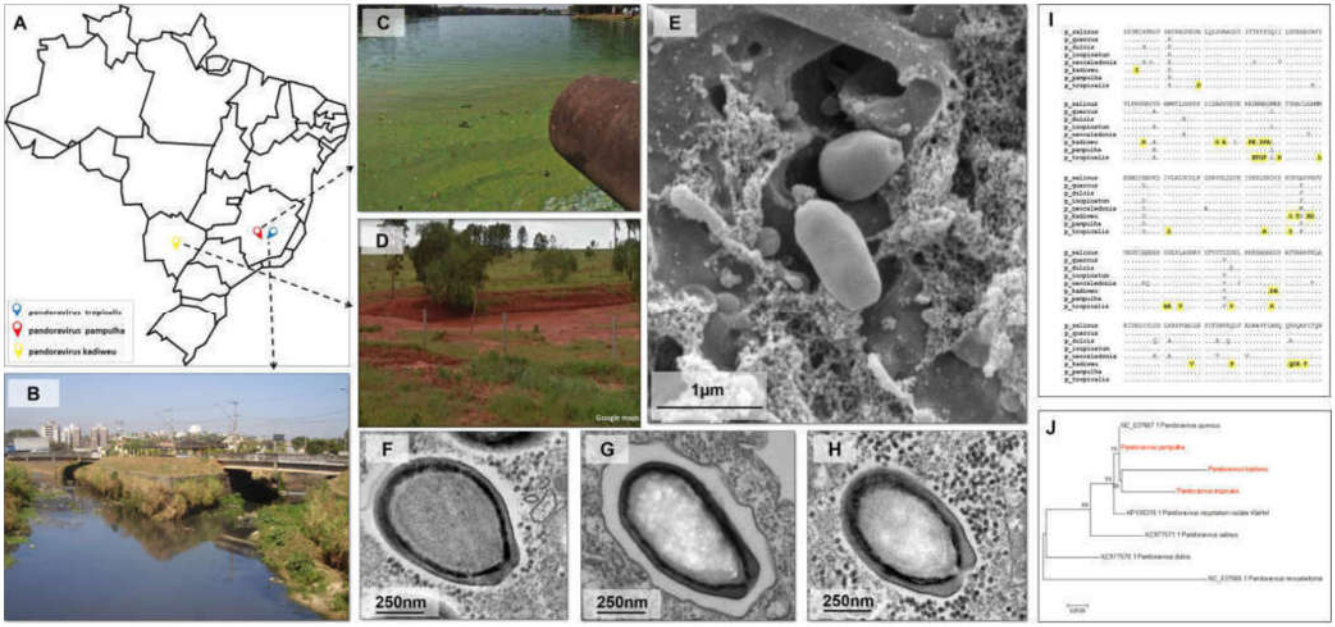


FIG 1 Sites of collection and electron microscopy and phylogenetic analysis of the pandoravirus isolated in this work. (A) Map of Brazil showing where the samples were collected for the isolation of pandoraviruses. (B to D) Representative pictures from the areas of collection: Bom Jesus Creek (B), Mergulhão Creek (C), and the city of Bonito (D). (E) *P. tropicalis* particles were analyzed using scanning electron microscopy at 24 h.p.i. and an MOI of 0.01. (F, G, and H) Transmission electron microscopy (24 h.p.i./MOI 0.01) for the viral particles corresponding to the isolates of *P. pampulha*, *P. tropicalis*, and *P. kadiweu*, respectively. (I) Alignment of the sequences, showing that *P. kadiweu* and *P. tropicalis* represent strains of pandoraviruses with many exclusive polymorphisms (highlighted in yellow), compared to the sequences of other isolated pandoraviruses. (J) Maximum likelihood tree constructed using predicted sequence of 251 amino acids of a DNA polymerase B subunit in different isolates of pandoraviruses. The giant viruses isolated in this work are highlighted in red.

sequenced a fragment of the DNA polymerase subunit B gene. The analysis of predicted amino acid sequences revealed that all of the isolates were different from each other. In addition, we observed that *P. tropicalis* and *P. kadiweu* present unique amino acid substitutions (Fig. 1I). The sequence of *P. pampulha* was more similar to that of *P. quercus* (Fig. 1I). These results reveal the diversity among our isolates and other pandoravirus isolates, and future genomic studies will determine whether *P. tropicalis* and *P. kadiweu* represent new clades among pandoraviruses (Fig. 1J). To date, there have been no rules or parameters available to establish a new clade belonging to the hypothetical family Pandoraviridae. The electron microscopy images obtained for these isolates were used to assemble a collection of more than 200 images. This data set allowed us to perform a comprehensive analysis of the replication cycle of these viruses.

Pandoraviruses are phagocytosed and replicate in large and electro-lucent viral factories. As demonstrated by work published by Legendre et al. in 2018 (14), the first steps involving the replication cycle of pandoraviruses seem to be similar for all these viruses, independently of the virus isolate analyzed. We observed that the amphora-shaped viral particles enter into acanthamoeba cells, likely by phagocytosis, which occurred within 30 min of infection (Fig. 2A and B). The particles were then transported to the interior of the amoebal cytoplasm, being carried inside phagosomes (Fig. 2C to E). This structure then seems to become fused with lysosome-like organelles, which, upon releasing their content inside the phagosome, stimulate the uncoating of the pandoravirus particles (Fig. 2C to F).

The viral factories (VFs) of the three analyzed isolates were wide, and electron-lucent areas occupied approximately 1/3 of the amoeba cytoplasm, containing viral particles in different stages of morphogenesis (Fig. 3). The VFs of the pandoraviruses seem to have been homogeneous and were not clearly limited by any cell component. Interestingly, we observed recruitment of mitochondria to regions inside and around the VFs (Fig. 3) and membranes were also recruited to regions inside the VFs (Fig. 4). In

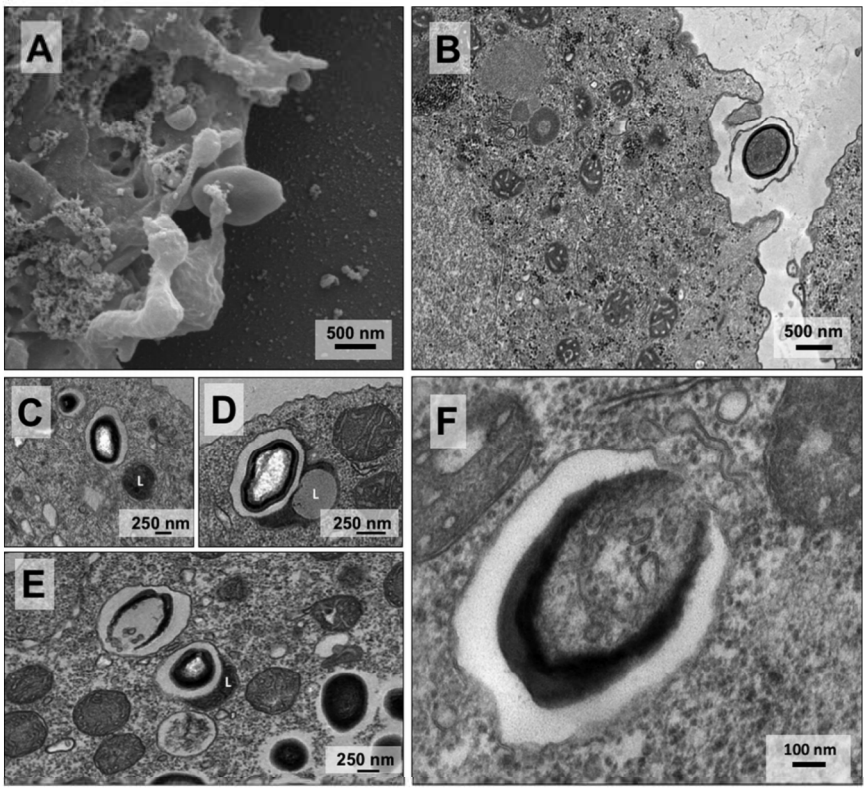


FIG 2 Initial steps of the pandoravirus replication cycle inside the amoebal host. (A and B) Scanning electron microscopy (A) and transmission electron microscopy (B) images show pandoravirus particles entering *Acanthamoeba castellanii* cells, likely as a consequence of phagocytosis. (C) The amoebas project pseudopods involving the viral particles and internalize them into vesicle-like structures known as phagosomes. (D and E) The phagosome then fuses with another component resembling a lysosome-like structure that, upon releasing their combined content, stimulates the uncoating of the pandoravirus particles (F). Although we used representative images in this figure, all the described steps were observed for all three isolated pandoraviruses. L, lysosome-like organelles; panels A and B, *Pandoravirus tropicalis*; panels C to F, *Pandoravirus kadiweu*.

addition, it is possible to observe an intense accumulation of structures that resemble lysosomal vesicles near the VFs (Fig. 3, orange arrows).

We also analyzed the appearance of the nuclear and nucleolar structures during the time course of infection of the three pandoravirus isolates (Fig. 5). The nuclear and nucleolar structures, appearing in the typical manner, were promptly observed both by transmission electron microscopy (TEM) and by Hemacolor staining in uninfected

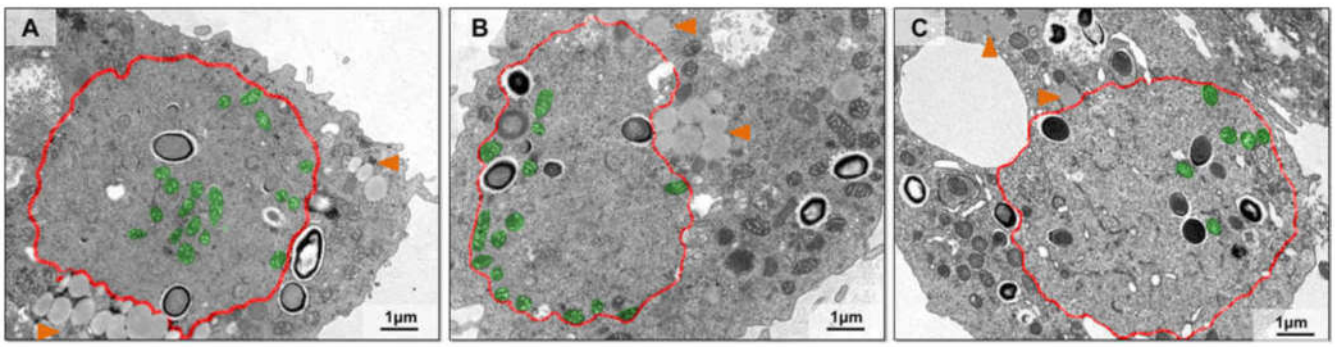


FIG 3 Characterization of pandoravirus viral factories. Viral factories of (A) *Pandoravirus tropicalis*, (B) *P. pampulha*, and (C) *P. kadiweu* were observed by transmission electron microscopy. The region of the viral factories is highlighted in red, the mitochondria present in the interior of the viral factories are highlighted in green, and the lysosomes are pointed out by orange arrowheads.

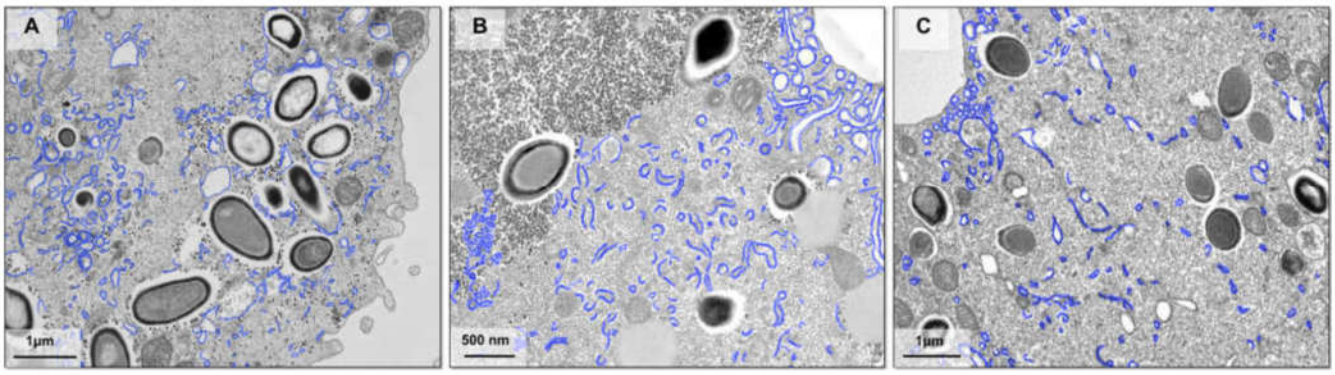


FIG 4 Membranes recruited inside pandoravirus viral factories. (A) Transmission electron microscopy of *P. tropicalis* viral factories. (B) Transmission electron microscopy of *P. pampulha* viral factories. (C) Transmission electron microscopy of *P. kadiweu* viral factories. The membranes recruited inside the viral factories are highlighted in blue.

acanthamoeba cells (Fig. 5A). As expected, the same was observed during viral entry (Fig. 5B). However, the nucleolar structure was no longer visible when the pandoraviruses' early VFs appeared, although we were still able to visualize the amoeba nucleus with its membrane (Fig. 5C). At late infection, the nuclear structure was no longer

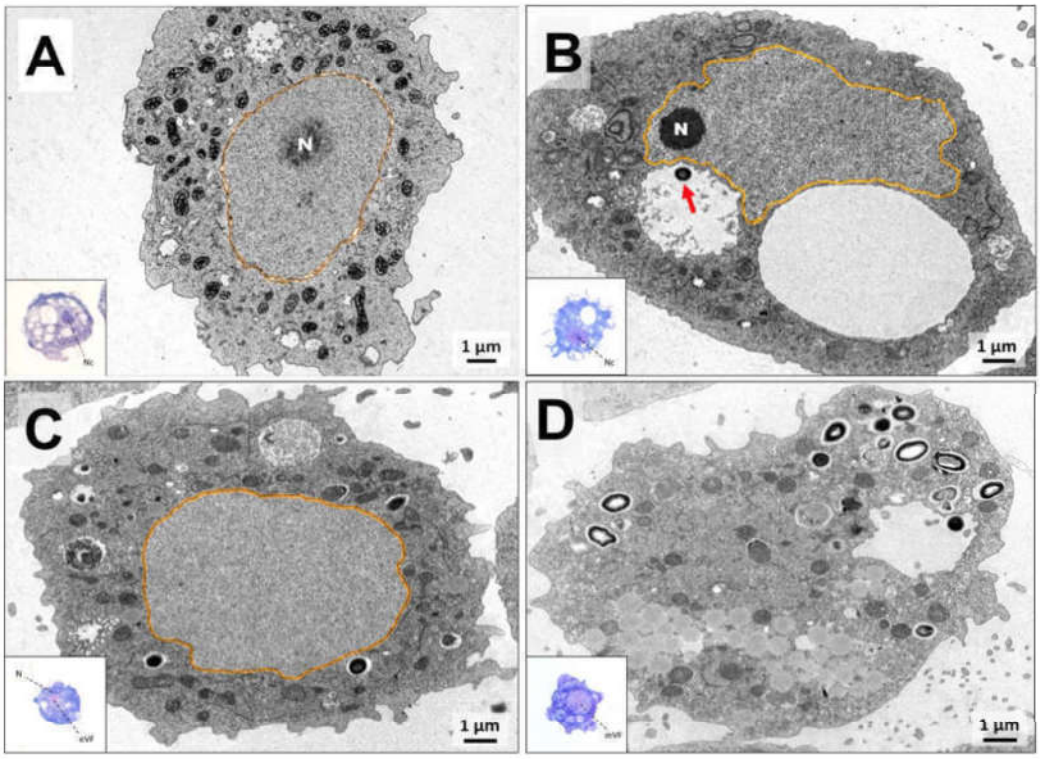


FIG 5 The *Acanthamoeba castellanii* cell nucleus becomes disorganized and loses its natural shape during the course of pandoravirus infection. (A) Transmission electron microscopy image showing a noninfected *Acanthamoeba castellanii* cell and how its nucleus is normally organized in this situation; it occupies about 2/3 of the cellular area, and it is delimited by a double-membrane layer known as the nuclear envelope (digitally highlighted in orange). The image at lower left represents the same conditions but visualized on a light microscope with Hemacolor staining. The nucleolus is observed as a dark spot surrounded by a bright area that represents the nucleus. (B) Image representing the amoeba observed just after the first steps of the pandoravirus replication cycle, as the virus (red arrow) is still harboring inside the amoebal phagosome. The nucleus does not yet seem to have suffered any modification at this stage. (C) At between h 3 and h 6 of infection, it seems that the nucleolus starts to be absent, as shown by one of the several images of transmission electron microscopy analyzed in this work. At lower left, the Hemacolor staining also shows the beginning of the appearance of the early viral factory. (D) The later steps of viral replication lead to the formation of the mature viral factory, marked by a bright area, easily recognizable in the images with Hemacolor staining. N, nucleus; Nc, nucleolus; eVF, early viral factory; mVF, mature viral factory.

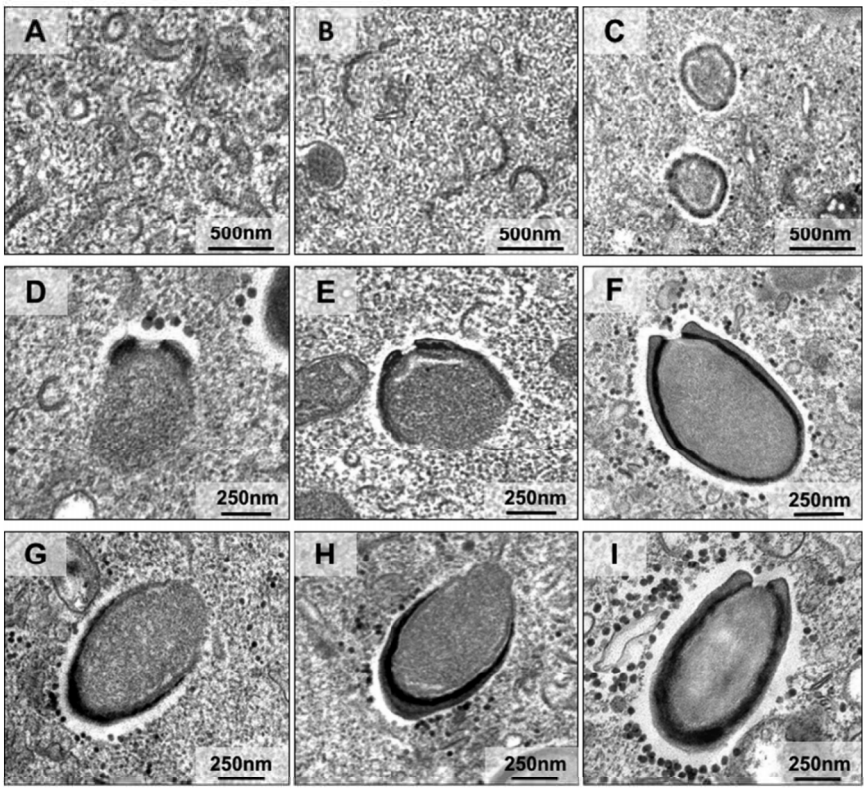


FIG 6 Morphogenesis of pandoravirus particles. Transmission electron microscopy images show stages of pandoravirus particle formation. (A to C) Crescent-like structures with different sizes, inside the viral factory, growing in thickness and complexity. (D to F) Particles being formed from the ostiolo-like apex. (G to I) Particles being formed from the end opposite the ostiolo-like apex. We used representative images of *P. tropicalis*, *P. pampulha*, and *P. kadiweu* in this panel; all the described steps were observed for the three isolates.

visible also, and the VFs occupied a substantial region in the cytoplasm (Fig. 5D). This process was observed during the replication cycle of the three isolates.

Morphogenesis dynamics of pandoravirus particles. After analysis of dozens of TEM images of asynchronous cycles of the isolated pandoraviruses, we noticed that the capsids of the pandoraviruses appeared to be formed from electron-dense semicircular structures observed in the middle of the VF (Fig. 6A). These structures appeared to become thicker and more electron dense as the cycle continued and to function as crescent-shaped precursors (Fig. 6B and C). The crescent-shaped precursors underwent a thickening of the apparent layer, followed by filling of the internal contents of the particles. As the adjacent portions of the capsids formed, the internal content of the particle continued to be filled simultaneously (Fig. 6D to I). As the particle enlarged, the capsid became more electron dense until closure of the total capsid, which at that stage was already filled with the particle's internal contents (Fig. 6I). After careful analysis of several images of our three isolates, we observed that particle morphogenesis/assembly could apparently start either at the ostiole-like apex or at the opposite end (Fig. 6D to I).

Pandoravirus particles are released by exocytosis and cell lysis. By studying the infection cycles of the new pandoravirus isolates, we made a curious observation. Analyses that have been done under a light microscope revealed that at early times of infection (until 6 h postinfection [h.p.i.]), these viruses could already be detected in the supernatant of infected cells, even at time points when the host cells had not yet undergone lysis. We then hypothesized that the pandoravirus particles could have started their release from the host by exocytosis, as suggested for some pandoravirus isolates (14). The analyses of TEM images of the new isolates revealed intense mem-

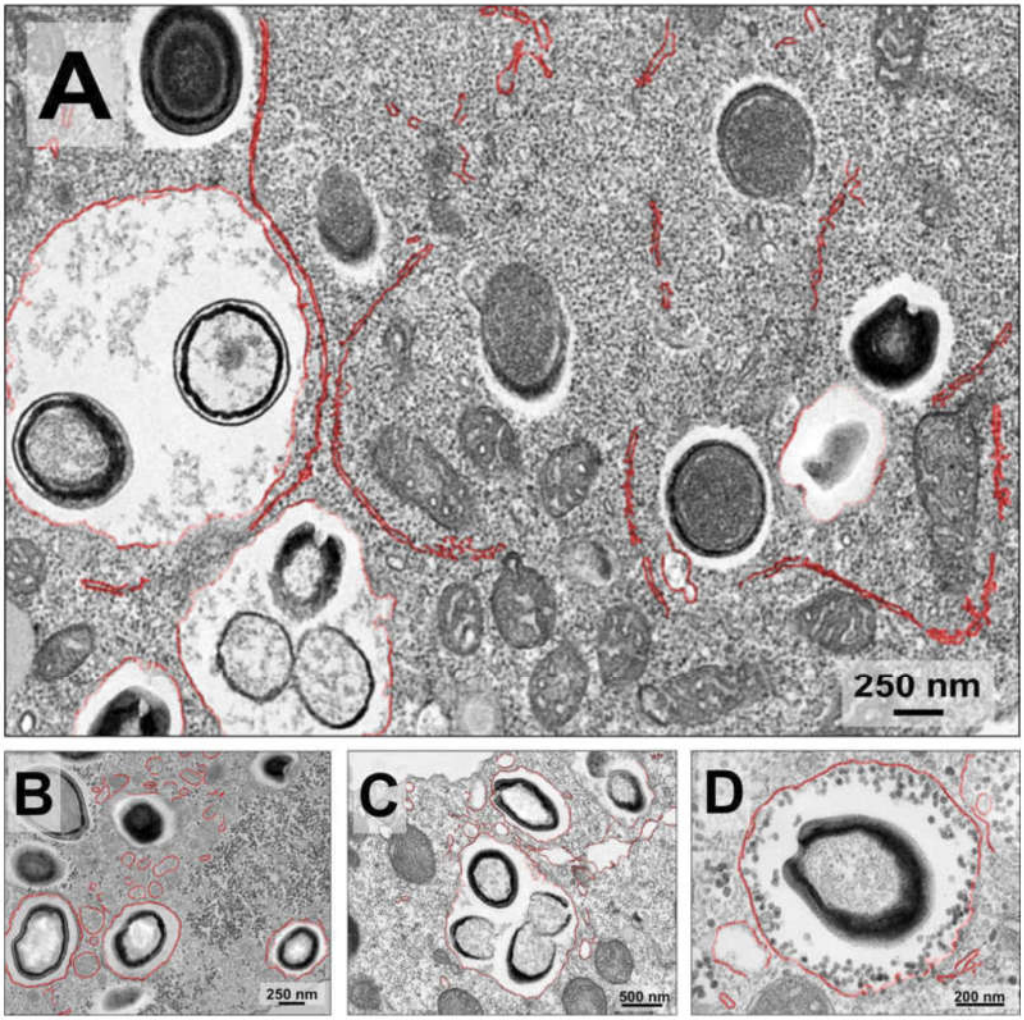


FIG 7 Transmission electron microscopy images showing pandoravirus particles being packaged into exosomes. (A) The late steps in pandoravirus replication are marked by intense membrane trafficking in the cytoplasm of the amoebal host (highlighted in red). This event is easily observed around the viral factory where the viral morphogenesis occurs. (B to D) Then, at around h 6 to h 9 postinfection, these double-membrane layers start to surround isolated or grouped viral particles, suggesting the beginning of exocytosis.

brane traffic close to just-formed particles, in the periphery of the VF (Fig. 7A). Interestingly, many particles were then wrapped inside such membranes, forming exosomes containing various amounts of viral particles of different sizes (Fig. 7). These exosome-containing particles then seemed to migrate to the periphery of the infected cell, fusing with the host cell cytoplasmic membrane and releasing the particles to the extracellular environment (Fig. 8).

To experimentally confirm that pandoraviruses can be released by exocytosis, we counted the acanthamoeba cells and the number of pandoraviruses particles in the supernatant through the viral cycle (multiplicity of infection [MOI] of 10). With this set of data, we analyzed whether the increase in the number of viral particles in the supernatant through the viral cycle could be observed before cell lysis was induced by viral infection, which would indicate that these viruses were being released by exocytosis at early times of infection. We observed that *P. tropicalis* caused the lysis of infected amoebas at 12 h.p.i., while no significant decrease in cell numbers was observed for cells infected with *P. pampulha* and *P. kadiweu* until 24 h.p.i. This indicates differences in the time postinfection at which each pandoravirus can induce host lysis (Fig. 9A to C). Cell lysis induced by *P. pampulha* and *P. kadiweu* was observed at 48 h.p.i.

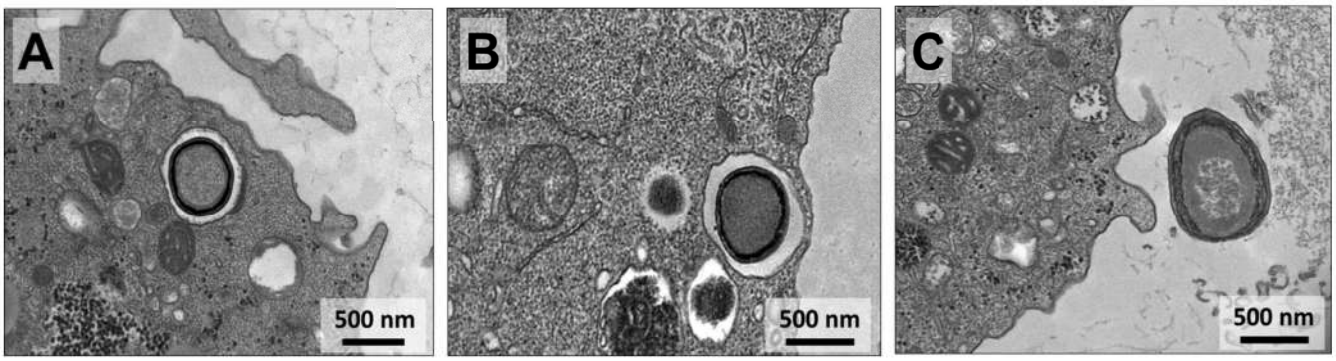


FIG 8 Transmission electron microscopy images demonstrating sequential steps of pandoravirus particle exocytosis. The images demonstrate that in late stages of infection the particles of pandoravirus start being packaged inside vesicles (A and B), becoming closer to the cytoplasmic membrane of the host cell and being released to the external milieu (C).

(data not shown). Interestingly, we observed an increase in the level of viral particles released in the supernatant from 6 h.p.i. for the three pandoravirus isolates, indicating that exocytosis might indeed contribute to particle release (Fig. 9D to F).

Aiming to evaluate the impact of membrane traffic inhibition in pandoravirus exocytosis, we pretreated infected amoebas with brefeldin A (BFA) (a membrane-trafficking inhibitor). Viral particles were counted at 12 h.p.i. for *P. pampulha* and *P. kadiweu* and at 6 h.p.i. for *P. tropicalis*. These time points were selected for each pandoravirus isolate based on the experiments last described above (Fig. 9A to F), whose results indicated the moment when the particles were undergoing exocytosis and the cells were not undergoing lysis. It was observed that acanthamoeba cultures treated with brefeldin A showed a reduction in the number of particles released for *P. pampulha* and *P. kadiweu* viruses (Fig. 9G and H). Curiously, the same was not observed for *P. tropicalis* (Fig. 9I). Future studies are needed to clarify why *P. tropicalis* can cause lysis of cells earlier than *P. kadiweu* and *P. pampulha* and why its exocytosis does not seem to be affected by brefeldin A treatment.

DISCUSSION

Giant virus prospective studies have revealed an outstanding universe of viral diversity (3, 4, 6, 16–20). Metagenomic studies have indicated the presence of a giant virus gene set in all continents (21–25). Some representatives, such as the mimivirus, appear to be more abundant and ubiquitous, containing hundreds of isolates already reported (21–26). Pandoravirus-like sequences were also detected in metagenomic data from environmental samples (22, 27, 28) as well as from insects, simian bushmeat, and human plasma (23–25, 27). Despite this, the amount of pandoravirus isolates is still limited (4, 6, 8, 11–13). Therefore, the isolation of new pandoraviruses contributes to the understanding of their biology, diversity, and distribution. The analyses of the isolates obtained in this work add important information characterizing the steps in the pandoravirus replication cycle.

It was hypothesized that pandoraviruses enter amoebas by phagocytosis (14). Our data for *P. tropicalis*, *P. pampulha*, and *P. kadiweu* reinforce this previous observation, as particles can be seen inside large vesicles in the amoebal cytoplasm within 30 min postinfection (Fig. 2C to E). Korn and Weisman demonstrated in 1967 that only particles larger than 500 nm can trigger phagocytosis in *Acanthamoeba*, a condition so far fulfilled by pandoravirus particles (29). Our images clearly demonstrate the induction of pseudopod formation when amoebas were kept in contact with pandoravirus particles (Fig. 2A and B). Despite this evidence, the possibility of pandoravirus particles entering amoebas by macropinocytosis could not be overruled, since this pathway also forms endosomes larger than 1 μm (30). However, the involvement of macropinocytosis in virion entry is a rare phenomenon in the literature (30). After entry of amoebas and release of the inner virion content (Fig. 2F), a short eclipse phase and an increase in

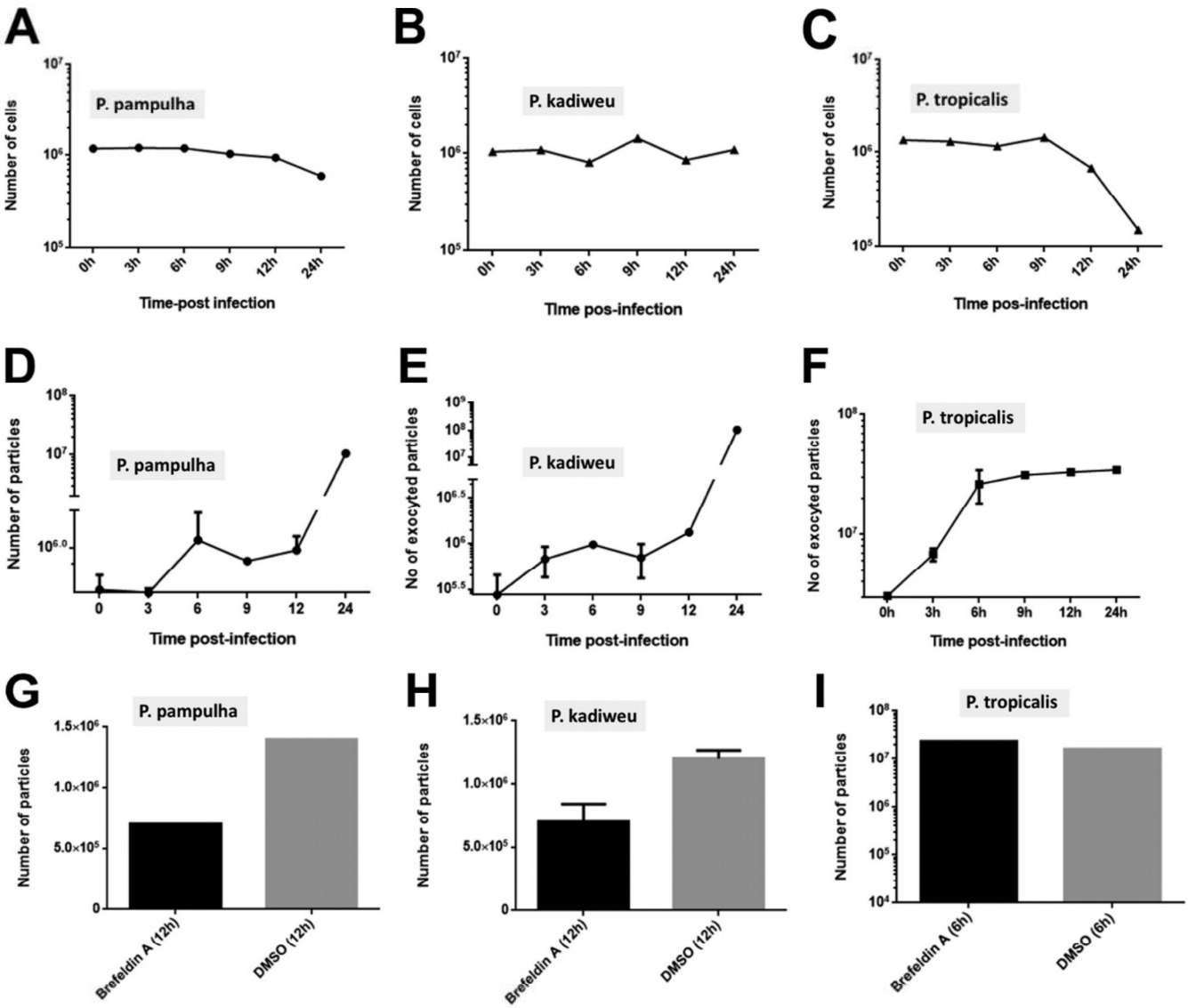


FIG 9 Analysis of the time course of infection for the pandoravirus isolates in *Acanthamoeba castellanii* cultures. The course of infection of *P. tropicalis*, *P. pampulha*, and *P. kadiweu* was established. (A to C) First, we observed the kinetics for the diminishing of the number of amoebas during the replication cycles of *P. pampulha* (A), *P. kadiweu* (B), and *P. tropicalis* (C) analyzed by cell counting. (D to F) Then, the number of viral particles present in the supernatant of these infected cells was also observed for *P. pampulha* (D), *P. kadiweu* (E), and *P. tropicalis* (F), at different time points. After setting a time point at which we observed an increase of more than 1 log of virus particles in the supernatant but without observing cellular lysis, the amoebal cultures were treated with an inhibitor of membrane trafficking (brefeldin A). (G to I) The cells were then infected with *P. pampulha* (G), *P. kadiweu* (H), and *P. tropicalis* (I) to check the influence of brefeldin A in the viral release. The number of exocyted viral particles in supernatant was counted. DMSO, dimethyl sulfoxide.

growing pandoravirus VFs occur (Fig. 3 and 5). As for the distribution of cellular organelles, the presence of many mitochondria inside and near the VF was seen (Fig. 3), along with the polarization of structures that resemble lysosomal vesicles in the vicinity of the VF (Fig. 3), as previously reported for the cedratvirus (31). The presence of mitochondria in this region could be related to the process of energy acquisition during viral replication optimization (32). Lysosome polarization, corresponding to the presence of vacuole-like structures occupying large portions of the host cell, might be related to a cellular response to infection, such as autophagy, as suggested previously in cedratvirus (31, 32). In addition, we observed gradual nucleolar and nuclear degradation throughout the pandoravirus replication cycle (Fig. 4), as observed for other pandoraviruses (6).

The onset of particle morphogenesis seems to occur from electron-dense semicir-

cular structures, such as the crescents observed in the assembly of other viruses of the NCLDV group, including vaccinia virus, mimivirus, marseillevirus, African swine fever virus, and cedratvirus (31, 33–36). The other morphogenesis steps of these viruses resemble those described for pandoraviruses, molliviruses, pithovirus, and other pandoraviruses, with the outer portion and the interior of the particles being assembled or "knitted" simultaneously (4, 6, 37). However, in addition to what was previously suggested, we observed that the viral particles seem to be assembled from both ends and not just from the ostiole-like apex (Fig. 6D to I) (4). We believe that studies about pandoraviruses assembly dynamics need more attention, since analyses limited to a few TEM sections could lead to misinterpretations. The similarities observed with respect to VF organization and the morphogenesis of pandoravirus and other viruses of the NCLDV group reinforce the previously suggested idea that these viruses could share an ancestor (38–41).

In a recent study, different strains of pandoravirus were seen to be initially released during the viral replication cycle by exocytosis processes (14). *Pandoravirus quercus*, *P. neocaledonia*, and *P. macleodensis* complete their entire replication cycle by between 8 to 12 h, starting the viral particles' exocytosis in about 8 h.p.i. and finishing their replicative cycle with lysis of the amoebal host cells, releasing hundreds of virions in the supernatant. Although the particles of the pandoravirus isolates analyzed here were seen to produce and were exocytosed from the cell as fast as those described by Legendre et al. (14) (in about 6 to 12 h.p.i.), the lysis of cells was observed at the late times of infection (12 to 48 h) (Fig. 9A to F). These results not only reinforce the hypothesis that pandoraviruses can explore different pathways for the progeny release but also demonstrate biological differences among viral isolates.

Many aspects of the replication cycle of these viruses still need to be clarified. This work provides new data and reveals new questions that future studies, especially at the molecular level, could answer. Prospective studies may also contribute in this regard by revealing novel members within the NCLDV group and improving understanding of the diversity, evolution, and biology of these complex viruses.

MATERIALS AND METHODS

Viral isolation, stock production, and titration. Three different pandoravirus isolates were used in this work. Two were coisolated with mimivirus from sewage samples collected in streams in the Pampulha region, Belo Horizonte, Brazil, in previous work and named *P. pampulha* and *P. tropicalis* (12). The other pandoravirus was isolated in the present work from water samples collected in Bonito, Mato Grosso do Sul, Brazil, and named *P. kadiweu*. For viral isolation, we used *A. castellanii* (ATCC 30234) as previously described (12). In order to produce the viral stocks, *A. castellanii* cells were grown in cell culture flasks and infected with 500 μl of isolates. After observation of typical cytopathic effect (cell rounding and lysis), the flask content was collected and the viruses were titrated. The titer was obtained by the endpoint method (42) and expressed as the number of 50% tissue culture infective doses ($TCID_{50}$) per milliliter. Viral stocks were kept at −70°C until use in further experiments.

Sequencing, alignment, and phylogeny. A fragment of the DNA polymerase B subunit gene (from position 473404 to position 474507—reference *Pandoravirus quercus*) was amplified (5′GCCCTCAAGCG GGGCCGCATG3′ and 5′CATCCACTGGGTGATCGGCGCCT3′) and sequenced, in both orientations and in duplicate, using an automated capillary sequencer (MegaBACE sequencer; GE Healthcare, Buckinghamshire, United Kingdom). For the phylogenetic tree, the resulting predicted amino acid (aa) sequences (251 aa) were aligned with previously published sequences obtained from GenBank using ClustalW in MEGA 7.0 software. This gene is highly conserved among pandoravirus strains and has been used in other studies (6, 12). The tree was constructed using the maximum likelihood method and a bootstrap value of 1,000.

Amoebal and viral particle counting. Initial electron microscopy analyses raised the hypothesis that pandoravirus could be released by exocytosis. In that way, two experiments were coupled and performed in duplicate that involved (i) the counting of intact amoebas throughout the viral replication cycle and (ii) the counting of pandoravirus particles that were being released in the supernatant at the same time points of infection. *A. castellanii* cells were infected in 25-cm^2 culture flasks (Kasvi, Brazil) with *P. tropicalis*, *P. pampulha*, and *P. kadiweu* isolates using an MOI of 10, and analyses were carried out at the infection times of 0 h, 3 h, 6 h, 9 h, 12 h, and 24 h. The time point of 0 h corresponds to an adsorption step of 30 min after infection, when the monolayer of cells was washed with 1× phosphate-buffered saline (PBS) and the flasks were filled with 4 ml of peptone-yeast extract-glucose (PYG) medium. After each time point was reached, we separated 12 μl of the supernatant to count the number of pandoravirus particles released during infection. The particles were observed under light microscopy (OlympusBX41, Japan), at ×1,000 magnification, using a cell counting chamber (Kcell Olen Kasvi, Brazil). In parallel, 12 μl of whole content presented in the flasks (including cells) was used to count the number of intact amoeba cells

observed at the different time points of the viral replication cycle. The same procedure was used to count the eukaryotic cells but at a magnification of ×400.

Brefeldin assays. The impact of brefeldin A (BFA) treatment on the pandoravirus replication cycle was verified. For this, 10^6 *A. castellanii* cells implanted in 25-cm^2 culture flasks were infected with the pandoravirus isolates at an MOI of 10 and treated with 10 μg/ml of BFA. We analyzed two different infection periods, 6 h.p.i. for *P. tropicalis* and 12 h.p.i. for *P. pampulha* and *P. kadiweu*. These periods correspond to the replication cycle stages before cell lysis for each virus occurs. The assays were performed in duplicate. Pandoravirus particles were counted using light microscopy as described above.

Hemacolor staining. *A. castellanii* cells were infected with isolates at an MOI of 10 and collected at 0 h.p.i., 3 h.p.i., 6 h.p.i., 9 h.p.i., 12 h.p.i., and 24 h.p.i. Then, 10 μl of the collected suspension was spread on a histological slide and were fixed with methanol. The VFs and viral particles were observed after Hemacolor staining (Renylab, Brazil), according to the manufacturer's recommendations. Slides were then analyzed under an optical microscope (OlympusBX41, Japan) at ×1,000 magnification.

Transmission electron microscopy. *A. castellanii* cells were infected at an MOI of 0.01 as described in the previous section and fixed at various times postinfection with 2.5% glutaraldehyde in a 0.1 M sodium phosphate buffer for 1 h at room temperature. The amoebas were postfixed with 2% osmium tetroxide and embedded in Epon resin. Ultrathin sections were then analyzed using transmission electron microscopy (TEM; Spirit Biotwin FEI-120 kV) at the Center of Microscopy of Universidade Federal de Minas Gerais (UFMG).

Scanning electron microscopy. *A. castellanii* cells infected at an MOI of 0.01 were added to round glass coverslips covered 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 postfixed with 1.0% osmium tetroxide for 1 h at room temperature. After the second fixation, the samples were washed three times with 0.1 M cacodylate buffer and immersed in 0.1% tannic acid for 20 min. The samples were then washed in cacodylate buffer and dehydrated by serial passages in ethanol solutions at concentrations ranging from 35% to 100%. Samples were then subjected to critical point drying using CO_2, placed in stubs, and metalized with a 5-nm-particle-size gold layer. The analyses were completed using scanning electron microscopy (FEG Quanta 200 FEI) at the Center of Microscopy of UFMG.

Accession number(s). Sequences are available in GenBank under accession numbers MK131392 (*P. kadiweu*), MK131393 (*P. pampulha*), and MK131394 (*P. tropicalis*).

ACKNOWLEDGMENTS

We thank our colleagues from Gepvig and the Laboratório de Vírus for their excellent technical support. We also thank CNPq, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), FAPEMIG, and the Center of Microscopy of UFMG.

E.G.K. and J.S.A. are CNPq researchers. E.G.K. and J.S.A. are members of a CAPES-COFECUB project.

We declare that we have no conflicts of interest.

REFERENCES

1. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie JM, Raoult D. 2003. A giant virus in amoebae. Science 299:2033. https://doi.org/10.1126/science.1081867.
2. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie JM. 2004. The 1.2-megabase genome sequence of Mimivirus. Science 306:1344–1350. https://doi.org/10.1126/science.1101485.
3. 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 U S A 106:21848–21853. https://doi.org/10.1073/pnas.0911354106.
4. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, Lescot M, Poirot O, Bertaux L, Bruley C, Coute Y, Rivkina E, Abergel C, Claverie JM. 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci U S A 111:4274–4279. https://doi.org/10.1073/pnas.1320670111.
5. Rodrigues RAL, Andreani J, Andrade A, Machado TB, Abdi S, Levasseur A, Abrahao JS, La Scola B. 13 June 2018. Morphologic and genomic analyses of new isolates reveal a second lineage of cedratviruses. J Virol https://doi.org/10.1128/JVI.00372-18.
6. Philippe N, Legendre M, Doutre G, Coute 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.
7. Scheid P, Zoller L, Pressmar S, Richard G, Michel R. 2008. An extraordinary endocytobiont in Acanthamoeba sp. *isolated from a patient with* keratitis. Parasitol Res 102:945–950. https://doi.org/10.1007/s00436-007-0858-3.
8. Scheid P, Balczun C, Schaub GA. 2014. Some secrets are revealed: parasitic keratitis amoebae as vectors of the scarcely described pandoraviruses to humans. Parasitol Res 113:3759–3764. https://doi.org/10.1007/s00436-014-4041-3.
9. Scheid P. 2016. A strange endocytobiont revealed as largest virus. Curr Opin Microbiol 31:58–62. https://doi.org/10.1016/j.mib.2016.02.005.
10. Antwerpen MH, Georgi E, Zoeller L, Woelfel R, Stoecker K, Scheid P. 2015. Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. Genome Announc 3:e00136-15. https://doi.org/10.1128/genomeA.00136-15.
11. Dornas FP, Khalil JY, Pagnier I, Raoult D, Abrahao J, La Scola B. 2015. Isolation of new Brazilian giant viruses from environmental samples using a panel of protozoa. Front Microbiol 6:1086. https://doi.org/10.3389/fmicb.2015.01086.
12. Andrade ACDSP, Arantes TS, Rodrigues RAL, Machado TB, Dornas FP, Landell MF, Furst C, Borges LGA, Dutra LAL, Almeida G, Trindade G. d S, Bergier I, Abrahão W, Borges IA, Cortines JR, de Oliveira DB, Kroon EG, Abrahão JS. 2018. Ubiquitous giants: a plethora of giant viruses found in Brazil and Antarctica. Virol J 15:22. https://doi.org/10.1186/s12985-018-0930-x.
13. Aherfi S, Andreani J, Baptiste E, Oumessoum A, Dornas FP, Andrade A, Chabriere E, Abrahao J, Levasseur A, Raoult D, La Scola B, Colson P. 2018. A large open pangenome and a small core genome for giant pandoraviruses. Front Microbiol 9:1486. https://doi.org/10.3389/fmicb.2018.01486.
14. Legendre M, Fabre E, Poirot O, Jeudy S, Lartigue A, Alempic JM, Beucher

L, Philippe N, Bertaux L, Christo-Foroux E, Labadie K, Coute Y, Abergel C, Claverie JM. 2018. Diversity and evolution of the emerging Pandoraviridae family. Nat Commun 9:2285. https://doi.org/10.1038/s41467-018-04698-4.

15. Abergel C, Legendre M, Claverie JM. 2015. The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. FEMS Microbiol Rev 39:779–796. https://doi.org/10.1093/femsre/fuv037.
16. 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.
17. Andreani J, Aherfi S, Bou Khalil JY, Di Pinto F, Bitam I, Raoult D, Colson P, La Scola B. 2016. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. Viruses 8:e300. https://doi.org/10.3390/v8110300.
18. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, Raoult D, La Scola B. 2016. Kaumoebavirus, a new virus that clusters with faustoviruses and Asfarviridae. Viruses 8:E278. https://doi.org/10.3390/v8110278.
19. Schulz F, Yutin N, Ivanova NN, Ortega DR, Lee TK, Vierheilig J, Daims H, Horn M, Wagner M, Jensen GJ, Kyrpides NC, Koonin EV, Woyke T. 2017. Giant viruses with an expanded complement of translation system components. Science 356:82–85. https://doi.org/10.1126/science.aal4657.
20. Abrahao 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.
21. Hingamp P, Grimsley N, Acinas SG, Clerissi C, Subirana L, Poulain J, Ferrera I, Sarmento H, Villar E, Lima-Mendez G, Faust K, Sunagawa S, Claverie JM, Moreau H, Desdevises Y, Bork P, Raes J, de Vargas C, Karsenti E, Kandels-Lewis S, Jaillon O, Not F, Pesant S, Wincker P, Ogata H. 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. ISME J 7:1678–1695. https://doi.org/10.1038/ismej.2013.59.
22. Kerepesi C, Grolmusz V. 2017. The "Giant Virus Finder" discovers an abundance of giant viruses in the Antarctic dry valleys. Arch Virol 162:1671–1676. https://doi.org/10.1007/s00705-017-3286-4.
23. Atoni E, Wang Y, Karungu S, Waruhiu C, Zohaib A, Obanda V, Agwanda B, Mutua M, Xia H, Yuan Z. 2018. Metagenomic virome analysis of Culex mosquitoes from Kenya and China. Viruses 10:E30. https://doi.org/10.3390/v10010030.
24. Temmam S, Monteil-Bouchard S, Robert C, Pascalis H, Michelle C, Jardot P, Charrel R, Raoult D, Desnues C. 2015. Host-associated metagenomics: a guide to generating infectious RNA viromes. PLoS One 10:e0139810. https://doi.org/10.1371/journal.pone.0139810.
25. Temmam S, Davoust B, Chaber AL, Lignereux Y, Michelle C, Monteil-Bouchard S, Raoult D, Desnues C. 2017. Screening for viral pathogens in African simian bushmeat seized at a French airport. Transbound Emerg Dis 64:1159–1167. https://doi.org/10.1111/tbed.12481.
26. Aherfi S, Colson P, La Scola B, Raoult D. 2016. Giant viruses of amoebas: an update. Front Microbiol 7:349. https://doi.org/10.3389/fmicb.2016.00349.
27. Verneau J, Levasseur A, Raoult D, La Scola B, Colson P. 2016. MG-Digger: an automated pipeline to search for giant virus-related sequences in metagenomes. Front Microbiol 7:428. https://doi.org/10.3389/fmicb.2016.00428.
28. Brinkman NE, Villegas EN, Garland JL, Keely SP. 2018. Reducing inherent biases introduced during DNA viral metagenome analyses of municipal wastewater. PLoS One 13:e0195350. https://doi.org/10.1371/journal.pone.0195350.
29. Korn ED, Weisman RA. 1967. Phagocytosis of latex beads by Acanthamoeba. II. Electron microscopic study of the initial events. J Cell Biol 34:219–227. https://doi.org/10.1083/jcb.34.1.219.
30. Mercer J, Helenius A. 2008. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. Science 320:531–535. https://doi.org/10.1126/science.1155164.
31. Silva L, Andrade A, Dornas FP, Rodrigues RAL, Arantes T, Kroon EG, Bonjardim CA, Abrahao JS. 2018. Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. Sci Rep 8:4000. https://doi.org/10.1038/s41598-018-22398-3.
32. Novoa RR, Calderita G, Arranz R, Fontana J, Granzow H, Risco C. 2005. Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol Cell 97:147–172. https://doi.org/10.1042/BC20040058.
33. Suarez C, Andres G, Kolovou A, Hoppe S, Salas ML, Walther P, Krijnse Locker J. 2015. African swine fever virus assembles a single membrane derived from rupture of the endoplasmic reticulum. Cell Microbiol 17:1683–1698. https://doi.org/10.1111/cmi.12468.
34. Suarez C, Welsch S, Chlanda P, Hagen W, Hoppe S, Kolovou A, Pagnier I, Raoult D, Krijnse Locker J. 2013. Open membranes are the precursors for assembly of large DNA viruses. Cell Microbiol 15:1883–1895. https://doi.org/10.1111/cmi.12156.
35. Mutsafi Y, Zauberman N, Sabanay I, Minsky A. 2010. Vaccinia-like cytoplasmic replication of the giant Mimivirus. Proc Natl Acad Sci U S A 107:5978–5982. https://doi.org/10.1073/pnas.0912737107.
36. Andrade A, Rodrigues RAL, Oliveira GP, Andrade KR, Bonjardim CA, La Scola B, Kroon EG, Abrahao JS. 2017. Filling knowledge gaps for mimivirus entry, uncoating, and morphogenesis. J Virol 91:e01335-17. https://doi.org/10.1128/JVI.01335-17.
37. Legendre M, Lartigue A, Bertaux L, Jeudy S, Bartoli J, Lescot M, Alempic JM, Ramus C, Bruley C, Labadie K, Shmakova L, Rivkina E, Coute Y, Abergel C, Claverie JM. 2015. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. Proc Natl Acad Sci U S A 112:E5327. https://doi.org/10.1073/pnas.1510795112.
38. Filee J, Chandler M. 2010. Gene exchange and the origin of giant viruses. Intervirology 53:354–361. https://doi.org/10.1159/000312920.
39. 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.
40. 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.
41. Iyer LM, Balaji S, Koonin EV, Aravind L. 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res 117:156–184. https://doi.org/10.1016/j.virusres.2006.01.009.
42. Reed LJ, Muench H. 1938. A simple method of estimating fifty per cent endpoints. Am J Epidemiol 27:493–497. https://doi.org/10.1093/oxfordjournals.aje.a118408.

4. Artigos número 03 e 04:

4.1 Yaravirus: a novel 80-nm vírus infecting *Acanthamoeba castellanii*

(doi: https://doi.org/10.1073/pnas.2001637117)

4.2 *Yaraviridae*: a new family of virus infecting *Acanthamoeba castellanii*

(doi: 10.1007/s00705-021-05326-1)

À medida que novos isolados foram sendo descobertos, o estudo de virus gigantes de ameba se tornou bastante consolidado. Dentre diversas outras características, esses vírus foram observados como sendo predominantemente marcados por um grande tamanho de partícula, um genoma extenso que codifica de centenas a milhares de proteínas, e pela presença de uma vasta quantidade dessas proteínas apresentando funções que a princípio só eram observadas em organismos celulares. Nesse trabalho, descrevemos o isolamento e caracterização de um novo vírus de ameba, denominado de *Yaravirus brasiliense,* com características e origem bastante intrigantes quando comparado a outros isolados nesses protistas. O yaravírus é o primeiro isolado que infecta culturas de *Acanthamoeba castellanii* e que potencialmente pode não pertencer aos vírus do grupo NCLDV, atualmente classificados no filo *Nucleocytoviricota*. Os trabalhos com essa amostra se iniciaram em 2015, com a coleta e posteriores tentativas de prospecção de novos vírus gigantes em água coletada da Lagoa da Pampulha, Belo Horizonte. Análises iniciais de coloração negativa e MET nos mostraram a presença de um vírus com tamanho muito menor aos já observados. Posteriormente, por meio de diversos experimentos de imagem como MET, MEV, tomografia eletrônica e microscopia crio-eletrônica, conseguimos estabelecer uma descrição detalhada de diversas etapas do ciclo de multiplicação do yaravírus, que se assemelha em alguns pontos aos de outros de vírus de ameba, no entanto, também apresentando importantes variações que serão mais bem descritas a seguir. Sem sombra de dúvida, o ponto mais importante desse trabalho foram as análises do genoma completo desse vírus, que demostraram seu genoma de dupla fita de DNA que codifica para 74 proteínas, que cerca de 90% são ORFans e com nenhuma função já descrita previamente pela ciência. Essa grande proporção de ORFans torna bastante difícil entender a relação evolutiva dos yaravírus com outros grupos virais existentes devido à falta de sinal filogenético capaz de suportar essas análises. Para alguns genes, como no caso da ATPase e da bastante divergente proteína principal do capsídeo, análises filogenéticas conseguiram agrupar o yaravírus dentro do grupo dos nucleocitovírus, juntamente com sequências de metagenoma dos distantes Pleurochrysis sp. vírus. Por fim, tomando tudo em conjunto, as importantes diferenças de aspectos biológicos e genéticos entre os yaravírus e outros vírus de ameba nos permitiu a

criação de um novo gênero viral (“Yaravirus”) e uma nova família (“Yaraviridae”), dentro dos quais o *Yaravirus brasiliense* e outros vírus relacionados possam ser devidamente classificados. A seguir, apresentamos dois artigos referentes à presente tese, que descrevem a (i) descoberta e caracterização do yaravírus e (ii) a proposta de criação da família “Yaraviridae”, recentemente aprovada pelo ICTV.

Links para informações suplementares do artigo “Yaravirus: A novel 80-nm virus infecting *Acanthamoeba castellanii*”, PNAS 2020.

Material suplementar:
https://www.pnas.org/content/pnas/suppl/2020/06/26/2001637117.DCSupplemental/pnas.2001637117.sapp.pdf

Arquivo do *software* ARAGORN com informações sobre as tRNAs do yaravírus:
https://www.pnas.org/content/pnas/suppl/2020/06/26/2001637117.DCSupplemental/pnas.2001637117.sd01.pdf

Arquivo do *software* PHYRE2 com informações sobre as estruturas tridimensionais preditas de algumas proteínas do yaravírus:
https://www.pnas.org/content/pnas/suppl/2020/06/26/2001637117.DCSupplemental/pnas.2001637117.sd02.pdf

Vídeo suplementar 01:
https://movie-usa.glencoesoftware.com/video/10.1073/pnas.2001637117/video-1

Vídeo suplementar 02:
https://movie-usa.glencoesoftware.com/video/10.1073/pnas.2001637117/video-2

Vídeo suplementar 03:
https://movie-usa.glencoesoftware.com/video/10.1073/pnas.2001637117/video-3

Yaravirus: A novel 80-nm virus infecting *Acanthamoeba castellanii*

Paulo V. M. Boratto[a,b,1], Graziele P. Oliveira[a,b,1], Talita B. Machado[a], Ana Cláudia S. P. Andrade[a], Jean-Pierre Baudoin[b,c], Thomas Klose[d], Frederik Schulz[e], Saïd Azza[b,c], Philippe Decloquement[b,c], Eric Chabrière[b,c], Philippe Colson[b,c], Anthony Levasseur[b,c], Bernard La Scola[b,c,2], and Jônatas S. Abrahão[a,2]

[a]Laboratório de Vírus, Instituto de Ciências Biológicas, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil; [b]Microbes, Evolution, Phylogeny and Infection, Aix-Marseille Université UM63, Institut de Recherche pour le Développement 198, Assistance Publique–Hôpitaux de Marseille, 13005 Marseille, France; [c]Institut Hospitalo-Universitaire Méditerranée Infection, Faculté de Médecine, 13005 Marseille, France; [d]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907; and [e]Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Edited by James L. Van Etten, University of Nebraska-Lincoln, Lincoln, NE, and approved June 2, 2020 (received for review January 29, 2020)

Here we report the discovery of Yaravirus, a lineage of amoebal virus with a puzzling origin and evolution. Yaravirus presents 80-nm-sized particles and a 44,924-bp dsDNA genome encoding for 74 predicted proteins. Yaravirus genome annotation showed that none of its genes matched with sequences of known organisms at the nucleotide level; at the amino acid level, six predicted proteins had distant matches in the nr database. Complimentary prediction of three-dimensional structures indicated possible function of 17 proteins in total. Furthermore, we were not able to retrieve viral genomes closely related to Yaravirus in 8,535 publicly available metagenomes spanning diverse habitats around the globe. The Yaravirus genome also contained six types of tRNAs that did not match commonly used codons. Proteomics revealed that Yaravirus particles contain 26 viral proteins, one of which potentially representing a divergent major capsid protein (MCP) with a predicted double jelly-roll domain. Structure-guided phylogeny of MCP suggests that Yaravirus groups together with the MCPs of *Pleurochrysis* endemic viruses. Yaravirus expands our knowledge of the diversity of DNA viruses. The phylogenetic distance between Yaravirus and all other viruses highlights our still preliminary assessment of the genomic diversity of eukaryotic viruses, reinforcing the need for the isolation of new viruses of protists.

Yaravirus | ORFan | NCLDV | metagenomics | capsid

Viral evolution and classification have been subjects of an intense debate, especially after the discovery of giant viruses that infect protists (1–4). These viruses are predominantly characterized by the large size of their virions and genomes encoding hundreds to thousands of proteins, of which a large proportion currently remain without homologs in public sequence databases (5–10). These coding sequences are commonly referred as ORFans, and due to the lack of phylogenetic information, their origin and function still represent a mystery (11–14). Strikingly, the increasing number of available viral genomes demonstrates that there is a huge set and great diversity of genes without homologs in current databases, which needs to be further explored (11). Importantly, many amoebal virus ORFan genes have already been proven to be functional, being expressed and encoding for components of the viral particles (6, 15). However, the large set of ORFans makes it difficult to predict the biology of viruses discovered through cultivation-independent methods, such as metagenomics analysis, reinforcing the need for the complementary isolation and experimental characterization of new viruses. All currently known isolated amoebal viruses are related to nucleocytoplasmic large DNA viruses (NCLDVs) (16). This group comprises families of eukaryotic viruses (Poxviridae, Asfarviridae, Iridoviridae, Ascoviridae, Phycodnaviridae, Marseilleviridae, and Mimiviridae) as well as other amoebal virus lineages including pithoviruses, pandoraviruses, molliviruses, medusaviruses, pacmanviruses, faustoviruses, klosneuviruses, and others. NCDLVs have dsDNA genomes and were proposed to share a monophyletic origin based on criteria that include the sharing of a set of ancestral vertically inherited genes (17, 18). From this handful set of genes, a core gene cluster is found to be present in almost all members of the NCLDVs, being composed by five distinct genes, namely a DNA polymerase family B, a primase-helicase, a packaging ATPase, a transcription factor, and a major capsid protein (MCP) for which the double jelly-roll (DJR) fold constitutes the main protein architectural class (19, 20). Recently, the International Committee on Taxonomy of Viruses (ICTV) brought forward a proposal for megataxonomy of viruses (21). The DJR major capsid protein (MCP) supermodule of DNA viruses includes NCLDVs and other icosahedral viruses that infect prokaryotes and eukaryotes. In addition to the signature DJR-MCPs, the majority of these viruses also encode for additional single jelly-roll minor capsid proteins (e.g., penton proteins) and genome packaging ATPases of the FtsK-HerA superfamily.

According to this proposal, the evolutionary conservation of the three genes of the morphogenetic module in the DJR-MCP

Significance

Most of the known viruses of amoeba have been seen to share some features that eventually prompted authors to classify them into common evolutionary groups. Here we describe Yaravirus, an entity that could represent either the first isolated virus of *Acanthamoeba* spp. out of the group of NCLDVs or, in an alternative scenario, a distant and extremely reduced virus of this group. Contrary to what is observed in other isolated viruses of amoeba, Yaravirus does not have a large/giant particle or a complex genome, but at the same time carries a number of previously undescribed genes, including one encoding a divergent major capsid protein. Metagenomic approaches also testified for the rarity of Yaravirus in the environment.

Author contributions: F.S., P.D., E.C., P.C., A.L., B.L.S., and J.S.A. designed research; P.V.M.B., G.P.O., T.B.M., A.C.S.P.A., J.-P.B., T.K., F.S., S.A., P.D., P.C., A.L., B.L.S., and J.S.A. performed research; E.C., B.L.S., and J.S.A. contributed new reagents/analytic tools; P.V.M.B., G.P.O., J.-P.B., T.K., F.S., S.A., P.D., E.C., P.C., A.L., B.L.S., and J.S.A. analyzed data; and P.V.M.B., G.P.O., T.K., F.S., P.C., A.L., B.L.S., and J.S.A. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: Yaravirus genome data have been deposited at GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under accession number MT293574.

[1]P.V.M.B. and G.P.O. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: bernard.la-scola@univ-amu.fr or jonatas.abrahao@gmail.com.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2001637117/-/DCSupplemental.

First published June 29, 2020.

supermodule, to the exclusion of all other viruses, justifies the establishment of a realm named *Varidnaviria* (21). Although some NCLDVs, as pandoraviruses, seem to have lost the DJR-MCP gene, their genome harbors a large set of genes that support their classification into NCLDVs (and *Varidnaviria*, consequently). Here we describe the discovery of Yaravirus, an amoeba virus with a puzzling origin and evolution. This virus has a genome that mainly consists of a near full set of genes that are ORFans. Yaravirus could represent either the first isolated virus of *Acanthamoeba* spp. out of the group of NCLDVs, inaugurating a group belonging to *Varidnaviria*, or, in an alternative evolutive scenario, a distant and extremely reduced virus of this group. Viral particles have a size of 80 nm, escaping the concept of large and giant viruses. Thus, Yaravirus expands our knowledge about viral diversity and evolution.

Results

Yaravirus Isolation and Replication Cycle. A prospecting study was conducted by collecting samples of muddy water from creeks of an artificial urban lake called Pampulha, located at the city of Belo Horizonte, Brazil. Here, by using a protocol of direct inoculation of water samples on cultures of *Acanthamoeba castellanii* (Neff strain, ATCC 30010), we have managed to isolate an amoebal virus that we named *Yaravirus brasiliensis*, as a tribute to an important character (Yara, the mother of waters) of the mythological stories of the Tupi-Guarani indigenous tribes (22). Negative staining revealed the presence of small icosahedral particles on the supernatant of infected amoebal cells, measuring about 80 nm in diameter (Fig. 1*A*). Cryoelectron microscopy images of purified particles suggest that Yaravirus particles present two capsid shells, as previously described for *Faustovirus*, although future studies are needed to confirm this (23) (*SI Appendix*, Fig. S1).

At the beginning of infection in *A. castellanii*, Yaravirus particles are found attached to the outside part of the amoebal plasma membrane, suggesting the participation of a host receptor in order to internalize the virions (Fig. 1*B*, *SI Appendix*, Fig. S2, blue arrows, and Movie S1). The replication cycle is then followed by the incorporation of individual or grouped Yaravirus particles inside endocytic vesicles, which, in a later stage of infection, are found next to a region occupied by the nucleus (Fig. 1 *B–D* and *SI Appendix*, Fig. S2, red arrows). The viral factory then takes place and completely develops into its mature form, replacing the region formerly occupied by the cell nucleus and recruiting mitochondria around its boundaries, likely to optimize the availability of energy to construct the virions (Fig. 1*E*). The step corresponding to viral morphogenesis happens similarly as how it is observed for other viruses of amoeba. First, it starts by the appearance of small crescents in the electron-lucent region of the factory (Fig. 1*E* and 2*A*). Next, step by step, the virions gain an icosahedral symmetry by the sequential addition of more than one layer of protein or membranous components around its structure (Fig. 2*A*). The constructed virions, with a capsid still empty, start then to migrate to the periphery of the viral factory, where there is the accumulation of corpuscular electron-dense material (Figs. 1*E* and 2 *B* and *C*, *SI Appendix*, Fig. S3 *A–C*, and Movie S2). These enucleations are scattered throughout the periphery of the infected cell and seem to represent different regions or morphogenesis points where the final step for Yaravirus maturation occurs. In these regions, the capsid of Yaravirus is filled with electron-dense material and the virus is finally ready to be released (Fig. 2*C*, red arrow, and *SI Appendix*, Fig. S4). Sometimes it is also possible to observe several particles of Yaravirus being packed in the interior of vesicle-like structures, suggesting a potential release by exocytosis, as observed for other viruses of amoeba (24, 25) (Fig. 2*D* and Movie S3). Most of the viral shedding, however, still occurs by lysis of the amoebal cell, followed by the release of Yaravirus particles, which later reach the supernatant of the infected culture, or sometimes might get

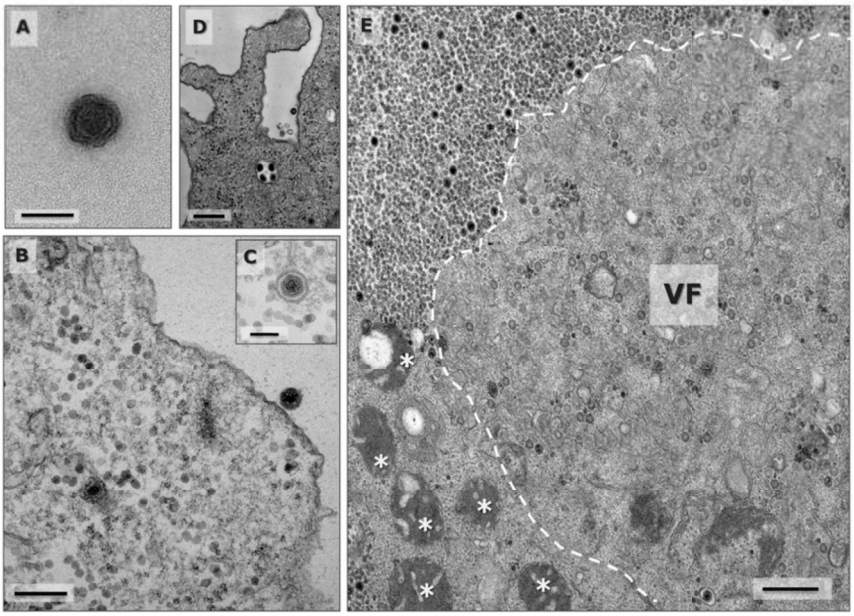


Fig. 1. Yaravirus particle and the beginning of the viral cycle. (*A*) Negative staining of an isolated Yaravirus virion. (Scale bar: 100 nm.) (*B*) Transmission electron microscopy (TEM) representing the beginning of the viral cycle, in which one particle is associated to the host cell membrane and the second one was already incorporated by the amoeba inside an endocytic vesicle. (Scale bar: 200 nm.) (*C*) Detailed image of an incorporated Yaravirus particle in the interior of an endocytic vesicle. (Scale bar: 100 nm.) (*D*) Viral uptake by the amoeba may occur individually but also in groups of particles, as observed in the micrograph. (Scale bar: 250 nm.) (*E*) The viral factory completely develops, occupying the nuclear region and recruiting mitochondria around it. Two different regions can be distinct: an electron-lucent region where the virions are assembled as empty shells and a second region formed by several electron-dense points where the genome is packaged inside the particles. (Scale bar: 500 nm.)

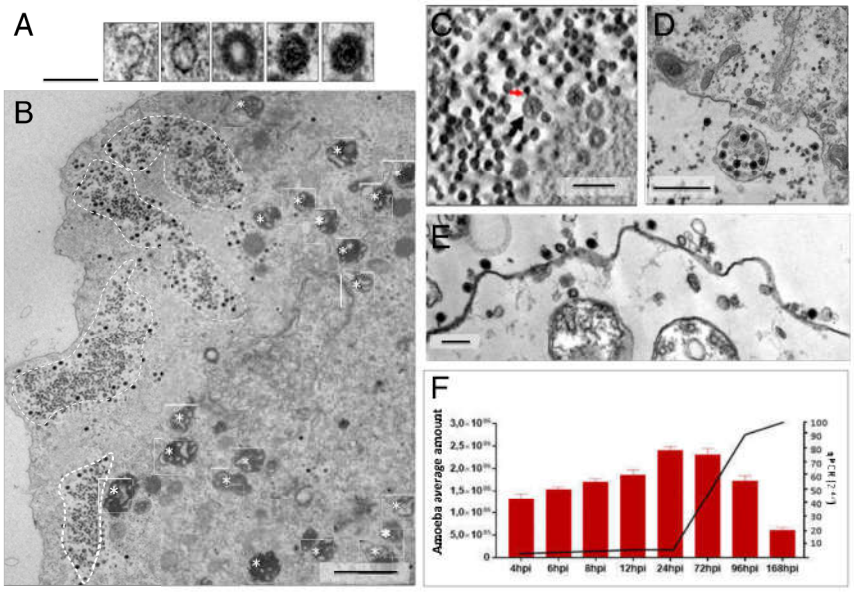


Fig. 2. Yaravirus morphogenesis and release. (*A*) The virions are assembled by the addition of more than one layer of protein or membranous components around its structure. (Scale bar: 70 nm.) (*B*) The particles then start to migrate to the periphery of the cell, where there is the presence of several electron-dense points that function as morphogenetic structures to package the DNA inside the Yaravirus particles (regions inside dashed lines). (Scale bar: 1,000 nm.) (*C*) Detailed image of the morphogenetic regions where the DNA (red arrow) is incorporated inside the Yaravirus virion (black arrow). (Scale bar: 150 nm.) (*D*) Sometimes, the final step of viral replication is marked by the particles being packaged inside vesicle-like structures, suggesting a potential release by exocytosis. (Scale bar: 500 nm.) (*E*) Most of the particles, however, are released by cellular lysis and have a high affinity to the membranes of cellular debris. (Scale bar: 150 nm.) (*F*) Graph comparing concomitantly the decrease of host cell numbers (red bars) with the increase of Yaravirus genome during the infection (black line). Replication of viral genome was measured by qPCR and calculated by delta-delta Ct.

attached to the debris of the cellular membranes (Fig. 2*E*). We have also evaluated Yaravirus replication by concomitantly investigating the decrease of the host cell numbers together with the increase of viral genome during infection. Interestingly, during the first hours of infection, the *A. castellanii* cultures seem to progressively grow until 24 h.p.i., showing a fastidious character for Yaravirus replication (Fig. 2*F*). The cells then start to suffer lysis induced by the virus only after 72 h.p.i. (Fig. 2*F*). On the same level, from 96 h.p.i. to 7 d post infection, there is no change of the detection levels of Yaravirus genome and the lysis seems to stop, and the remaining trophozoites turn to cysts.

Genome. Sequencing of the Yaravirus genome has shown the presence of a double-stranded DNA molecule with a length of 44,924 bp and harboring a total of 74 predicted genes (Fig. 3*A*). Although two of these genes (73 and 74) seemed to be truncated and do not start with codons belonging to the methionine amino acid, both were detected in the proteomic analysis (*Yaravirus* Proteomics). Despite a smaller genome than other viruses of amoeba, Yaravirus encodes for six tRNA genes: tRNA-Ser (gct), tRNA-Ser (tga), tRNA-Cys (gca), tRNA-Asn (gtt), tRNA-His (gtg), and tRNA-Ile (aat) (Fig. 3 *A* and *D* and *SI Appendix*, Tables S1 and S2 and Dataset S1). All of them are colocated on an intergenic region between genes 29 and 30 (Fig. 3 *A* and *D*). In contrast to tRNA genes in Tupanviruses, we did not observe a correlation between the Yaravirus tRNA isoacceptors and the codons most frequently used by the virus or its *A. castellanii* host (*SI Appendix*, Fig. S5). The genome has a GC content of 57.9%, which is one of the highest found in any amoebal virus discovered to date (*SI Appendix*, Fig. S6*A*). When analyzed gene by gene, Yaravirus has a spectrum of GC content that varies between 46% and 65% (*SI Appendix*, Table S3). The analysis of the intergenic regions of the genome (46% GC content) did not reveal any enriched sequence motifs that might indicate a conserved promoter, as opposed to what is observed in many other NCLDV members (26).

By considering only the portions of genome that are part of coding regions, Yaravirus also has a similar coding capacity as observed in other viruses when their genome was first annotated, ~90% (*SI Appendix*, Fig. S6*B*). Surprisingly, Yaravirus genome annotation showed that none of its genes matched with sequences of known organisms when we compared them at the nucleotide level. When we looked for homology at the amino acid levels, we found that only two predicted proteins had hits in the Pfam-A database and, in total, six had distant matches in the nr database. Therefore, considering the same criteria that have been used to analyze other giant viruses' genomes, about 90% ($n = 68$) of the Yaravirus predicted genes are ORFans. The six genes whose product has some homology with known protein sequences (Fig. 3*A* and Table 1) are homologous to fragments of proteins predicted to have different functions, such as an exonuclease/recombinase bacterial protein (gene 2; best hit, *Timonella senegalensis*), a hypothetical protein (gene 3; best hit, *A. castellanii*), a hypothetical protein (gene 28; best hit, *Acytostelium subglobosum* LB1, a dictyostelid), a packaging ATPase (gene 40; best hit, *Pleurochrysis* endemic virus), a conserved hypothetical protein (gene 46; best hit, *Melbournevirus*, a marseillevirus strain), and a bifunctional DNA primase/polymerase (gene 69; best hit, *Marinobacter* sp., Alphaproteobacteria; Table 1). Complimentary prediction of three-dimensional structures of these proteins indicated a potential function of 11 more genes (Table 2 and *Yaravirus Proteomics*). If we consider an additional 11 genes whose functions were predicted by structural analyses, the Yaravirus proportion of ORFans would be 80% ($n = 57$; still a high number, considering that, usually, in other studies involving giant viruses, methodologies based on protein structures are not used for this accounting) (27, 28). Phylogenetic analyses were then performed for genes 02, 40, 46, and 69 after aligning them with protein sequences of similar function belonging

to different members of the virosphere and to organisms of the three cellular domains of life. Other identified genes (Table 2) did not have enough genetic information to be included in a phylogenetic analysis. For analysis corresponding to the putative exonuclease/recombinase (gene 02), three major groups were observed to construct the morphology of the tree. Yaravirus was placed in one of those clades, clustering with some members of Eukarya, specifically with stony coral and insects (*SI Appendix*, Fig. S7). Analyses of gene 40 (virion packing ATPase) revealed that Yaravirus clustered in a polyphyletic branch, with members belonging to Mimiviridae family, bacteria (although many of these sequences seem to represent misclassified NCLDVs from metagenome-assembled genomes), and *Pleurochrysis* sp. endemic virus 1a and 2 (*SI Appendix*, Fig. S8). For the phylogenetic analysis corresponding to gene 46 (hypothetical protein conserved in *Marseillevirus*), Yaravirus was clustered with *Marseillevirus* strains (*SI Appendix*, Fig. S9). For the last tree, representing analysis for gene 69 (bifunctional DNA primase/polymerase), we have observed that Yaravirus was clustered with members of eukaryotes corresponding to the *Streblomastix* and *Phytophthora* groups (*SI Appendix*, Fig. S10). However, it should be noted that, in a previous study, the authors detected sequences of mimivirus genes among the *Phytophthora* parasitic strain INRA-310 genome (29). After all those analyses, it is important to note that, although Yaravirus has some genes with representatives in the genome of other organisms, their homology with orthologs is very low (25.24 to 44.12%), highlighting that Yaravirus genome content is essentially divergent among the other members of the virosphere (Table 1).

In order to detect sequences related to Yaravirus, we surveyed 8,535 publicly available metagenomes in the IMG/M database that have been generated from samples from diverse habitats across our planet (30). We discovered distant homologs of the Yaravirus ATPase (NCVOG0249) with an amino acid homology of up to 33.9% in the metagenomic data, while the closest homolog in the NCBI nr database was that of *Pleurochrysis* sp. endemic virus 1a, with 33.1%. In a phylogenetic tree of the viral ATPase, the Yaravrus branched within the Mimiviridae as part of a highly supported clade made up by its distant metagenomic relatives and *Pleurochrysis* sp. endemic viruses (Fig. 4 and *SI Appendix*, Fig. S11). In contrast to known members of the Mimiviridae, viral contigs and viral genomes in this clade featured a high GC content, with up to 62%. Adding sequences of polinton and virophage ATPases to this dataset resulted in a slightly altered topology, but the position of Yaravirus remained stable, not grouping with these additional elements (*SI Appendix*, Fig. S12). We also searched for proteins similar to the Yaravirus putative MCP but were not able to retrieve closely related sequences in the metagenomic data. In parallel, we used 18 hidden Markov models to detect MCPs in 235 NCLDV reference genomes (plus contigs of three *Pleurochrysis* endemic viruses). After dereplicating the hits by clustering at 90% sequence similarity using accurate mode in cd-hit, we prepared a structure-guided alignment with Expresso in T-Coffee (using PDB structures), and also a conventional alignment with mafft-ginsi (–unalignlevel 0.8–allowshift). Both alignments returned surprisingly good results for Yaravirus MCP, and then phylogenetic trees were calculated with iq-tree LG+F+R8. Yaravirus MCP was shown to group together with the MCPs of *Pleurochrysis* endemic viruses in a well-supported clade; however, in contrast to the ATPase tree, they were affiliated with Phycodnaviridae (chlorella viruses in particular) and not with the Mimiviridae (Fig. 5 and *SI Appendix*, Figs. S13 and S14 and Supplementary Files).

Finally, trying to investigate a potential relationship between Yaravirus and different members of the *Pleurochrysis* sp. endemic virus group, we have looked for protein similarities between these organisms. Comparisons were made using the BLASTp database and have suggested some ortholog candidates, but with a low percentage of protein identity (around 24 to 33%). For *Pleurochrysis* sp. endemic virus 1a and *Pleurochrysis* sp. endemic virus 1b (Genbank accession codes KY131436 and KY203336, respectively), there is some similarity with genes 28, 40, and 41 of Yaravirus. For *Pleurochrysis* sp. endemic virus 2

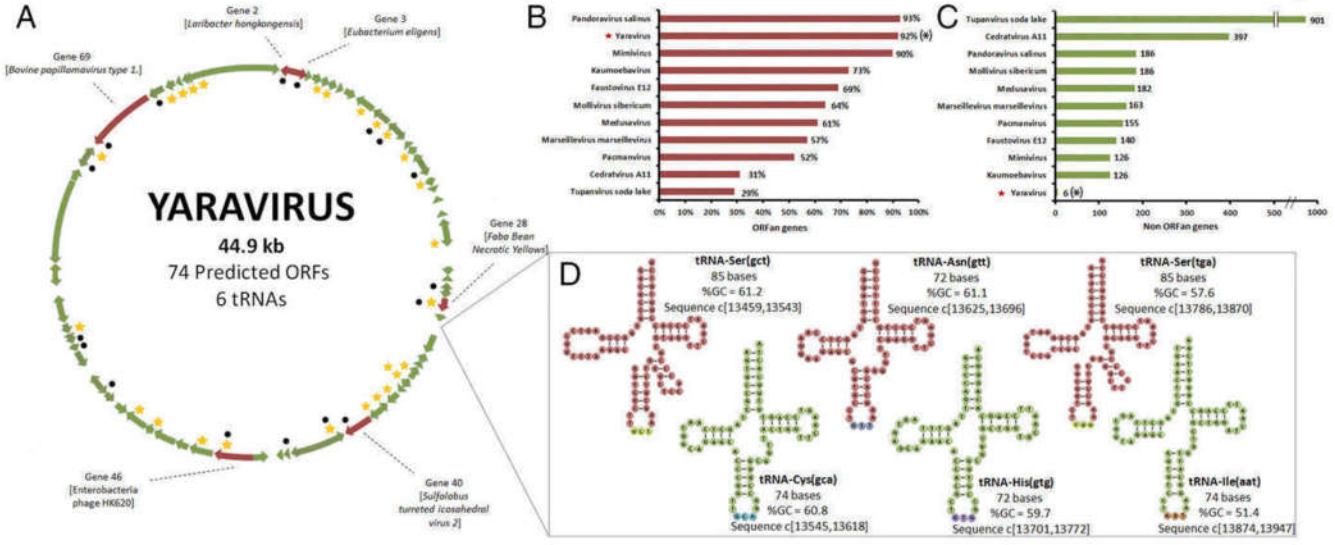


Fig. 3. Yaravirus genome features. (*A*) Circular representation of Yaravirus genome highlighting the predicted ORFs (arrows). Red arrows represent ORFs predicted by analyses of similarity of amino acid sequences with information regarding their best hits. Black dots indicate ORFs encoding predicted proteins whose functions were suggested by HHPred (structural analyses). Yellow stars indicate proteins found in virion proteomics. (*B*) The percentage of ORFan genes among the complete genome of different viruses of amoeba is represented by the graph with red scale bars. (*C*) The graph with greenish scale bars represents the absolute number of genes with homologs in databases (non-ORFan genes) for each of the same amoebal viruses previously analyzed. (*D*) All of the six Yaravirus predicted tRNAs, as well as their corresponding sequences, are pictured with information about their anticodon (in parentheses), their nucleotide length, the percentage of GC content, and the position in the intergenic regions of genes 29 and 30. Regarding the asterisk, note that the number of ORFan/non-ORFan genes represented in *C* and *D* do not take into account the structural annotation made by the HHpred servers, as for most of the amoebal viruses represented in the graphs.

Table 1. Yaravirus genes with similarity on current databases and their best-hits (BLASTp)

Yaravirus gene ID	Best hit	Total score	Query cover, %	E-value	Identity, %	Annotation of best BLASTp hit
2	*Timonella senegalensis* WP_019148817.1	67.0	87	1e-09	27.88	Exonuclease/recombinase
3	*A. castellanii* XP_004339080.1	51.6	80	1e-06	44.12	Hypothetical protein
28	*Acytostelium subglobosum* LB1 XP_012747655.1	57.8	85	4e-07	27.59	Hypothetical protein
40	*Pleurochrysis* sp. endemic virus 1a AUD57256.1	121	66	4e-28	33.05	Virion packaging ATPase
46	Melbournevirus YP_009094634.1	181	88	5e-47	32.63	Hypothetical protein conserved in marseilleviruses
69	*Marinobacter* sp. MAB50943.1	106	35	8e-20	25.24	Bifunctional DNA primase/polymerase

(Genbank accession code KY346835), two genes are related to gene 28 in Yaravirus and two others are similar to genes 40 and 02. Finally, for *Pleurochrysis* sp. endemic virus unk (accession code KY203337), we have found similarity with gene 69 of Yaravirus.

Yaravirus Proteomics. As aforementioned, most Yaravirus proteins had no detectable homologs in public databases. This peculiarity prompted us to have a closer look at the proteins responsible to form the mature particles of Yaravirus. Proteomics revealed a total of 26 viral proteins present in purified particles (*SI Appendix*, Table S4). We then analyzed the predicted three-dimensional structures of those 26 proteins by using three platforms for domain comparison, the HHpred, the Phyre2, and the Swiss-model tools (30–37). Only five sequences (genes 11, 12, 28, 41, and 46) were observed to have structural features similar to known proteins (Table 2 and Dataset 2). That means that about 80% of its virion proteome consists of ORFans. It is important to mention that the same approach (in silico structure prediction) has been used in parallel to evaluate all of the 74 predicted genes on the genome of Yaravirus, resulting in total in the discovery of 17 gene-encoded products with structural resemblance to other proteins in public databases (Table 2). Proteomics data revealed that the most abundant proteins in the viral particles corresponded to genes 41, 46, and 51 (from most to least abundant; *SI Appendix*, Table S4). While, for the third highest expressed protein, we were not able to find any structural candidates with known biological function, for sequences represented by genes 41 and 46, we observed fragments of protein resembling the three-dimensional structure of the capsid of other viruses (Dataset S2 and Table 2). With a confidence of 97%, a relevant portion (65%) of gene 41 was found to have a structural similarity with the double jelly-roll domain of the MCP of the *Paramecium bursaria* Chlorella virus type 1 (Dataset S2). Gene 46-encoded predicted protein was found to have structural similarity with bacterial secreted protein pcsB and tail needle protein (Table 2 and Dataset S2), a portion composed of a long alpha-helix. The function of protein encoded by gene 46 remains to be investigated. Therefore, we were not able to convincingly find any minor capsid protein. It is also important to note that sequences represented by gene 46 are the same described earlier to be highly conserved in marseilleviruses and in medusaviruses (Table 1). Finally, the last two proteins observed in the proteome that had matches with structural deposits in public databases are represented by genes 11 and 12, the first one predicted by HHpred to encode for an adiponectin and the second one for a protein called cerebellin-1 (Table 2).

Discussion

In recent years, amoebal large and giant viruses have frequently been found around the world (5–10, 22, 24, 27, 38–41). Here, we describe *Yaravirus brasiliensis*, an 80-nm-sized virus with a genome containing a notable proportion of genes that have never been observed before. Using standard protocols, our very first genetic analysis was unable to find any recognizable sequences of capsid or other classical viral genes in Yaravirus. This

Table 2. Annotation of Yaravirus proteins based on the predicted tridimensional structure of the proteins coded by the virus

Gene	Best hit	Functional prediction	Probability	E-value
2	*Laribacter hongkongensis*	Exonuclease	99.98	2.5e-30
3	*Eubacterium eligens*	Uncharacterized protein	94.49	0.88
11	***Homo sapiens***	**Adiponetctin**	**95.76**	**0.096**
12	***Rattus norvegicus***	**Cerebellin-1; synapse protein**	**97.54**	**5,00e-03**
17	*Rattus norvegicus*	prkc apoptosis wt1 regulator protein	95.86	0.17
26	*Homo sapiens*	Retinoblastoma-binding protein 6	94.71	0.026
28	**Faba Bean Necrotic Yellows**	**Replication-associated protein; endonuclease**	**95.07**	**0.11**
40	*Sulfolobus* turreted icosahedral virus 2	Genome packaging NTPase B204; FtsK-HerA superfamily	99.67	7.9e-15
41	**Singapore grouper iridovirus**	**Major capsid protein**	**98.89**	**3.1e-8**
43	*Arabidopsis thaliana*	Transcription factor HY5	96.6	0.016
46	**Enterobacteria phage HK620**	**DNA stabilization protein; tail needle, viral genome-ejection**	**86.58**	**0.033**
54	*Haloarcula marismortui*	Ribosome 50S	90.84	0.81
57	*H. sapiens*	BEN domain-containing protein 3	86.72	0.43
58	*Oryctolagus cuniculus*	Potential copper-transporting atpase	91.25	0.62
67	*Xenopus laevis*	DNA-(apurinic or apyrimidinic site) lyase	92.31	0.45
69	Bovine papillomavirus type 1	Replication protein E1/DNA Complex; DNA helicase, AAA+, ATPase	99.18	4.5e-10
70	*Escherichia coli*	Holliday junction resolvase	99.85	9.2e-20

Note: The bold genes represent proteins observed on the viral proteomics.

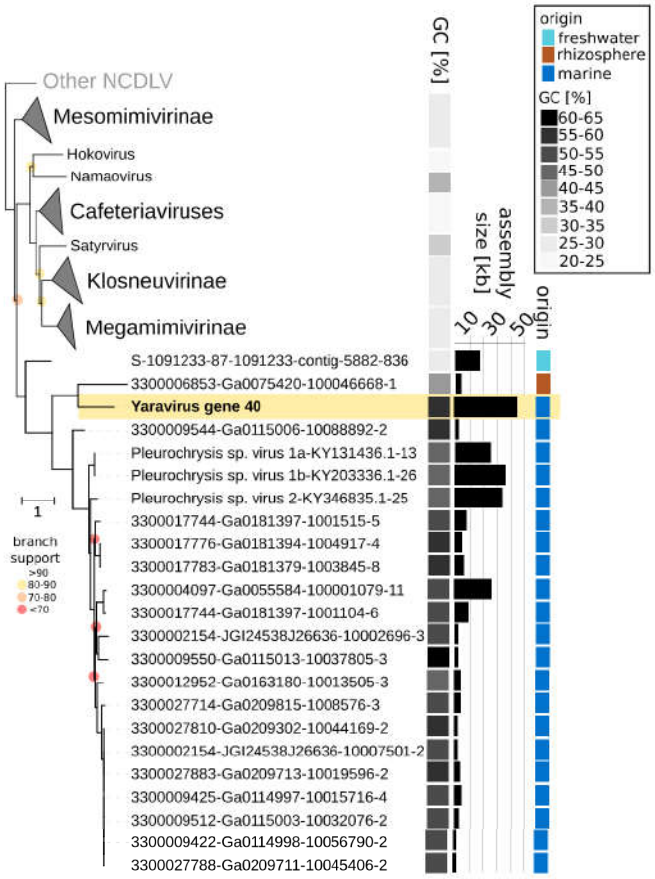


Fig. 4. Phylogenetic position of Yaravirus and related viral sequences in the Mimiviridae based on the viral ATPase (NCVOG0249). The Yaravirus ATPase is highlighted in yellow. Branch support is indicated as colored circles for support values of 90 or below. The tree is rooted at the Poxviruses. (Scale bar: substitutions per site.) GC content of viral genomes and contigs containing NCVOG0249 is shown together with the average GC content of collapsed clades. In addition, environmental origin and assembly sizes of Yaravirus and related viral contigs and genomes are shown.

is a relevant feature to highlight the importance of studies related to the isolation of new viral samples, as, by following the current metagenomic protocols for viral detection, Yaravirus would not even be recognized as a viral agent (42, 43).

According to our knowledge, Yaravirus represents the first virus isolated in *Acanthamoeba* spp. that is potentially not part of the complex group of NCLDVs. Several characteristics unite previously discovered amoebal viruses: large-sized virions, genomes coding for hundreds to thousands of genes, and presumably a monophyletic origin that is reflected in the presence of a set of about 20 most likely vertically inherited genes (17, 18). None of these features are present in Yaravirus, and that makes it potentially the first isolate of a novel bona fide group of amoebal virus. Of course, we cannot exclude the possibility that Yaravirus may represent a reduced NCLDV, presenting highly divergent or even absent NCLDV hallmark proteins. Recently, a similar case was described for three small crustacean viruses (44). However, despite their reduced genome when compared to other members of the NCLDV, an important number of hallmark genes were shared with this group, differently as observed for Yaravirus (44). In this not less exciting scenario, supported by the analysis regarding the ATPase and MCP phylogenetic trees, Yaravirus would represent the to-date smallest member of the NCLDVs, both in particle and genome size. The presence of six copies of tRNAs in Yaravirus also impresses when analyzed by the perspective of a selective pressure forcing to maintain these genes in such a small genome when compared to larger viruses of amoeba. Even more interestingly, none of the isoacceptors related to the Yaravirus tRNAs correspond to codons of amino acids abundantly used by the virus or the amoeba. Considering the fastidious infection cycle of Yaravirus in *Acanthamoeba*, it is conceivable that, in nature, a different organism might act as the preferred host of Yaravirus. Some genes were found to be shared between Yaravirus and members of the *Pleurochrysis* endemic virus group; however, the low coverage supporting their putative orthologs make difficult a close relationship between these organisms.

Most members of the to-date isolated giant viruses of amoeba show a capsid specially composed by copies of an MCP related to the D13L of Vaccinia virus (15, 45). Pandoraviruses are an exception, as they seem to lack a protein shell to protect their genomes (28). Interestingly, even some of the amoeba hosts of these viruses may carry copies of MCP genes, suggesting possible horizontal gene transfer between virus and protist host (46, 47). By the structure-guided alignment and analysis of the constructed phylogenetic trees, Yaravirus does seem to share a common origin of MCP capsid with other NCLDVs, even though their sequences show an incredibly low similarity (48). Taken together, we can conclude that Yaravirus represents a divergent lineage of viruses isolated from *A. castellanii*. The large amount of unknown proteins encoded by Yaravirus reflects the variability existing in the viral world and the astonishing coding potential of new viral genomes yet to be discovered.

Methods

Origin of Samples and Viral Isolation. In 2017, searching to isolate novel variants of viruses infecting amoebas, we collected samples of muddy water from a creek of Lake Pampulha, an artificial lagoon located at the city of Belo Horizonte, Brazil (19 51 0.60S and 43 58 18.90W). As soon as they were collected, the samples were quickly taken to our lab and stored at 4°C until they were further processed. Following the protocol, 4×10^4 amoebas of the *A. castellanii* Neff strain (ATCC 30010) were seeded in each well of a 96-well plate, inoculating to each one a volume of around 100 μL of the collected samples, originally diluted 1:10 in PBS buffer. The plates were then

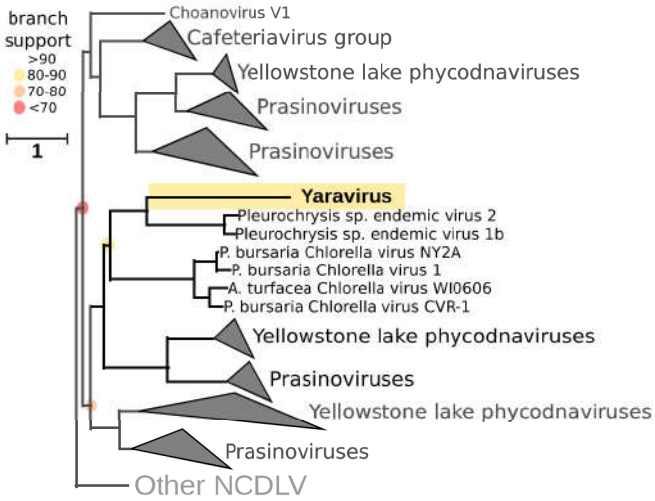


Fig. 5. Structure-guided phylogeny based on hmmsearch employed to identify the Yaravirus MCP in 235 NCDLV reference genomes using specific hidden Markov models. The alignments were made with Expresso in the software T-Coffee, using PDB structures. After, the phylogenetic trees were built using IQ-tree (v1.6.12; ref. 61) with LG+F+R8 based on the built-in model select feature (62) and 1,000 ultrafast bootstrap replicates (63). The Yaravirus MCP is highlighted in yellow. Branch support is indicated as colored circles for support values of 90 or below. The tree is unrooted. (Scale bar: substitutions per site.)

incubated for 7 d at 32°C and observed daily for the appearance of cytopathic effect, which may indicate a probable viral infection. All of the content from the wells was then collected and submitted to three processes of freezing and thawing and analysis of the possible isolates by negative staining technique. By the end, the collected content was submitted to another two blind passages in fresh cultures of amoeba, but this time in 25-cm^2 Nunc Cell Culture Treated Flasks with Filter Caps (Thermo Fisher Scientific) containing around 1 million amoebal cells. After viral isolation, all of the following experiments were made by infecting *A. castellanii* cells in a low multiplicity of infection (MOI), given the Yaravirus's fastidious replication cycle.

Transmission Electron Microscopy (TEM), TEM Tomography, Cryo-Electron Microscopy. For resin embedding and transmission electron microscopy (TEM), *A. castellanii* cells infected with Yaravirus were fixed at 20 h postinfection with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer. Cells were washed three times with a solution of 0.2 M saccharose in 0.1 M sodium cacodylate. Cells were postfixed for 1 h with 1% OsO_4 diluted in 0.2 M potassium hexa-cyanoferrate (III)/0.1 M sodium cacodylate. After washes with distilled water, cells were gradually dehydrated with ethanol by successive 10-min baths in 30, 50, 70, 96, 100, and 100% ethanol.

Substitution was achieved by successively placing the cells in 25, 50, and 75% Epon solutions for 15 min. Cells were placed for 1 h in 100% Epon solution and in fresh Epon 100% overnight at room temperature. Polymerization took place with cells in fresh 100% Epon for 48 h at 60°. Ultrathin 70- or 300-nm-thick sections were cut with a UC7 ultramicrotome (Leica) and placed on HR25 300 Mesh Copper/Rhodium grids (TAAB). Ultrathin sections were contrasted according to Reynolds (49). Electron micrographs were obtained on a Tecnai G20 TEM operated at 200 keV equipped with a 4,096 × 4,096-pixel resolution Eagle camera (FEI). For tomography, gold nanoparticles 10 nm in diameter (ref. 741957; Sigma-Aldrich) were deposited on both faces of the sections prior to contrasting. Tomography tilt series were acquired on the G20 Cryo TEM (FEI) with the Explore 3D software (FEI) for tilt ranges of 110° with 1° increments. The mean applied defocus was −2 μm. The magnification ranged between 3,500 and 29,000 with pixel sizes between 3.13 and 0.37 nm, respectively. The image size was 40,962 pixels. The tilt series were aligned using ETomo from the IMOD software package (University of Colorado) by cross-correlation (50). The tomograms were reconstructed using the weighted back-projection algorithm in ETomo from IMOD. The average thickness of the obtained tomograms was 268.40 ± 64 nm (n = 16). Fiji/ImageJ (NIH) was used for making tomography movies (51). For cryoelectron microscopy assays, the supernatant of infected cultures of *A. castellanii* was collected after 7 d postinfection and submitted to a first round of centrifugation at 1,500 × g for 10 min, looking to pellet the cell debris from the virus present on the supernatant. Next, the portion containing the Yaravirus was then submitted to a second round of centrifugation, and the virus was concentrated by ultracentrifugation at 100,000 × g for 2 h. The following steps were previously described by Klose et al. (23). Briefly, 3 μL of virus solution was placed on glow-discharged C-Flat 2/2 grids (EMS) and plunge-frozen into liquid ethane using a Gatan Cryoplunge 3. Samples were then imaged on a Talos F200C (ThermoFisher Scientific) equipped with a Ceta camera (ThermoFisher Scientific).

Genome Sequence and Analysis. The Yaravirus genome was sequenced two times by using the Illumina MiSeq platform (Illumina) with the paired-end application. The generated reads were then assembled de novo by using the software ABYSS and SPADES, with the resulting contigs ordered by the Python-based CONTIGuator.py software. After, gene predictions were made by using the GeneMarkS tool (52). The functional annotation for the Yaravirus predicted proteins was made through searches against the GenBank NCBI nonredundant protein sequence database (nr), considering as homologous proteins only the sequences that presented an e-value < 1×10^{-3}. The annotation was refined by comparing the in silico-predicted protein structures of Yaravirus with domains of proteins present in different databases using three platforms: the HHpred, the Phyre2, and the Swiss-model tools (30–37). It is important to note that, for the HHpred, we considered true-positive protein structures only matches that had a probability >80% and e-value ~1, as suggested by Söding and colleagues (37). For the qPCR assays, the increase in genome replication was assessed in cultures of *A. castellanii* cells infected by Yaravirus in different time points (H4, H6, H8, H12, H24, H72, H96, and H168), using primers which were constructed based on the sequence of the gene 69 of Yaravirus (primers, 5′TGCAGCAAGTCGGTCAAGAT3′ and 5′AACTTCCACATGCGAAACGC3′). Conditions used in the assay were previously described (53).

The amino acid and codon usage data were compared to those presented by *A. castellanii* and by different strains of amoebal viruses. For this, the sequences were downloaded from the NCBI database and analyzed by using the software Artemis 18.0.3. The percentage of GC content and GC skew have also been analyzed by using the same software. Transfer RNA (tRNA) sequences were identified using the ARAGORN tool. Phylogenetic analyses were performed for the six proteins of Yaravirus holding similarities with other organisms on the NCBI database (Table 1). By using the ClustalW tool in the Mega 10.0.5 software program, amino acid sequences of these Yaravirus proteins were previously aligned with the corresponding sequences of representatives of the virosphere and from other cellular organisms belonging to the three domains of life. The analysis involved 55 amino acid sequences for gene 40, 49 amino acid sequences for gene 02, 13 amino acid sequences for gene 46, and 34 amino acid sequences for gene 69. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All of the trees were constructed by using the maximum likelihood evolution method, with the JTT matrix-based model and a bootstrap of 1,000 replicates (54).

Yaravirus Proteomics. In order to identify the proteins that make up Yaravirus particles, thirty 75-cm^2 cell culture flasks (Nunc), containing 7×10^6 *A. castellanii* cells per flask, were infected with the isolated virus, and the cytopathic effect was followed up to 7 d.p.i. After severe amoebal lysis, the content was collected and submitted to a first round of centrifugation, at 1,500 × g for 10 min, looking to pellet the cell debris from the virus present on the supernatant. Then, this viral portion was submitted to a second round of centrifugation, and the virus was concentrated by ultracentrifugation at 100,000 × g for 2 h. To finish, viral pellet was then prepared for a two-dimensional gel electrophoresis and analysis by matrix-assisted laser desorption/ionization and liquid chromatography-tandem mass spectrometry as described before by Reteno and colleagues (55).

Metagenomic Survey. The Yaravirus ATPase (NCVOG0249) and the putative major capsid protein (MCP) were used to query 8,535 publicly available metagenomes in the IMG/M database (30) using diamond BLASTp (v0.9.25.126; ref. 56). Resulting protein hits with more than 30% query and subject coverage and an e-value of at least 1e-5 were extracted from the metagenomic data. In parallel, hmmsearch (version 3.1b2; hmmer.org) was employed to identify and extract ATPases (NCVOG0249) and MCPs (multiMCP) from 235 NCDLV reference genomes using specific hidden Markov models (https://bitbucket.org/berkeleylab/mtg-gv-exp/). Extracted proteins were then combined with the Yaravirus queries. To remove most redundant sequences, the MCP data set was clustered in cd-hit (57) at an amino acid similarity level of 90% using the accurate mode. ATPases were then aligned with MAFFT-linsi (v7.294b; ref. 58) and MCPs in Expresso (59) (most accurate mode in t-coffee v_13.41.0.28, alignment guided by PDB structures), and the resulting alignments trimmed with trimal (v1.4, -gt 0.1; ref. 60). Phylogenetic trees were built using IQ-tree (v1.6.12; ref. 61) with LG+F+R5 (ATPase) and LG+F+R8 (MCP) based on the built-in model select feature (62) and 1,000 ultrafast bootstrap replicates (63). Phylogenetic trees were visualized with iTol [v5.8 (61)] and ete3 (64).

Data Availability. Data for the Yaravirus genome have been deposited to GenBank under accession number MT293574.

ACKNOWLEDGMENTS. We thank our colleagues from IHU (Aix Marseille University) and from Laboratório de Vírus (Universidade Federal de Minas Gerais) for their assistance, especially Said Mougari, Issam Hasni, Lina Barrassi, Priscilla Jardot, Erna Kroon, Claudio Bonjardim, Paulo Ferreira, Giliane Trindade, and Betania Drumond. In addition, we thank the Méditerranée Infection Foundation, Centro de Microscopia da Universidade Federal de Minas Gerais (UFMG), Pro Reitorias de Pesquisa e Pós-graduação da UFMG, Programa de Pós-graduação em Microbiologia da UFMG, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FAPEMIG (Fundação de Amparo à Pesquisa do estado de Minas Gerais) for their financial support. J.S.A. is a CNPq researcher. B.L.S., J.S.A., P.C., P.V.M.B. and G.P.O. are members of a CAPES-Comitê Francês de Avaliação da Cooperação Universitária project. This work was supported by the Agence Nationale de la Recherche (ANR), including the 'Programme d'investissement d'avenir' under the reference Méditerranée Infection 10-1AHU-03 and European funding FEDER PRIMI. The work that was conducted by F.S. at the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract DE-AC02–05 CH11231. It made use of resources of the National Energy Research Scientific Computing Center, which is also supported by the DOE Office of Science under Contract DE-AC02–05CH11231.

1. J. Guglielmini, A. C. Woo, M. Krupovic, P. Forterre, M. Gaia, Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19585–19592 (2019).
2. E. V. Koonin, N. Yutin, Multiple evolutionary origins of giant viruses. *F1000 Res.* **7**, 1840 (2018).
3. P. Colson *et al.*, Ancestrality and mosaicism of giant viruses supporting the definition of the fourth TRUC of microbes. *Front. Microbiol.* **9**, 2668 (2018).
4. P. Colson, Y. Ominami, A. Hisada, B. La Scola, D. Raoult, Giant mimiviruses escape many canonical criteria of the virus definition. *Clin. Microbiol. Infect.* **25**, 147–154 (2019).
5. J. Abrahão *et al.*, Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **9**, 749 (2018).
6. M. Legendre *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.* **112**, E5327–E5335 (2015).
7. J. Andreani *et al.*, Pacmanvirus, a new giant icosahedral virus at the crossroads between Asfarviridae and Faustoviruses. *J. Virol.* **91**, e00212-17 (2017).
8. J. Andreani *et al.*, Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. *Viruses* **8**, 300 (2016).
9. L. H. Bajrai *et al.*, Kaumoebavirus, a new virus that clusters with faustoviruses and Asfarviridae. *Viruses* **8**, 278 (2016).
10. F. Schulz *et al.*, Giant virus diversity and host interactions through global metagenomics. *Nature* **578**, 432–436 (2020).
11. M. Boyer, G. Gimenez, M. Suzan-Monti, D. Raoult, Classification and determination of possible origins of ORFans through analysis of nucleocytoplasmic large DNA viruses. *Intervirology* **53**, 310–320 (2010).
12. N. Siew, D. Fischer, Twenty thousand ORFan microbial protein families for the biologist? *Structure* **11**, 7–9 (2003).
13. Y. Yin, D. Fischer, Identification and investigation of ORFans in the viral world. *BMC Genomics* **9**, 24 (2008).
14. N. Siew, D. Fischer, Unravelling the ORFan puzzle. *Comp. Funct. Genomics* **4**, 432–441 (2003).
15. P. Renesto *et al.*, Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products. *J. Virol.* **80**, 11678–11685 (2006).
16. N. A. Khan, *Acanthamoeba: Biology and Pathogenesis* (Caister Academic Press, ed. 2, 2015).
17. E. V. Koonin, N. Yutin, Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism. *Adv. Virus Res.* **103**, 167–202 (2019).
18. L. M. Iyer, S. Balaji, E. V. Koonin, L. Aravind, Evolutionary genomics of nucleocytoplasmic large DNA viruses. *Virus Res.* **117**, 156–184 (2006).
19. M. Krupovic, E. V. Koonin, Multiple origins of viral capsid proteins from cellular ancestors. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2401–E2410 (2017).
20. N. Yutin, Y. I. Wolf, D. Raoult, E. V. Koonin, Eukaryotic large nucleo-cytoplasmic DNA viruses: Clusters of orthologous genes and reconstruction of viral genome evolution. *Virol. J.* **6**, 223 (2009).
21. E. V. Koonin *et al.*, Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* **84**, e00061-19 (2020).
22. A. C. D. S. P. Andrade *et al.*, Ubiquitous giants: A plethora of giant viruses found in Brazil and Antarctica. *Virol. J.* **15**, 22 (2018).
23. T. Klose *et al.*, Structure of faustovirus, a large dsDNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6206–6211 (2016).
24. A. C. D. S. Pereira Andrade *et al.*, New isolates of pandoraviruses: Contribution to the study of replication cycle steps. *J. Virol.* **93**, e01942-18 (2019).
25. M. Legendre *et al.*, Diversity and evolution of the emerging Pandoraviridae family. *Nat. Commun.* **9**, 2285 (2018).
26. G. P. Oliveira *et al.*, Promoter motifs in NCLDVs: An evolutionary perspective. *Viruses* **9**, 16 (2017).
27. D. Raoult *et al.*, The 1.2-megabase genome sequence of Mimivirus. *Science* **306**, 1344–1350 (2004).
28. N. Philippe *et al.*, Pandoraviruses: Amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* **341**, 281–286 (2013).
29. V. Sharma, P. Colson, R. Giorgi, P. Pontarotti, D. Raoult, DNA-dependent RNA polymerase detects hidden giant viruses in published databanks. *Genome Biol. Evol.* **6**, 1603–1610 (2014).
30. I. A. Chen *et al.*, IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* **47**, D666–D677 (2019).
31. A. Waterhouse *et al.*, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
32. S. Bienert *et al.*, The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).
33. N. Guex, M. C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* **30** (suppl. 1), S162–S173 (2009).
34. P. Benkert, M. Biasini, T. Schwede, Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343–350 (2011).
35. M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, T. Schwede, Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 10480 (2017).
36. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
37. J. Soding, A. Biegert, A. N. Lupas, The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
38. M. Boyer *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.* **106**, 21848–21853 (2009).
39. P. V. Boratto *et al.*, Niemeyer virus: A new mimivirus group A isolate harboring a set of duplicated aminoacyl-tRNA synthetase genes. *Front. Microbiol.* **6**, 1256 (2015).
40. R. A. L. Rodrigues *et al.*, Morphologic and genomic analyses of new isolates reveal a second lineage of cedratviruses. *J. Virol.* **92**, e00372-18 (2018).
41. L. K. D. S. Silva *et al.*, Cedratvirus getuliensis replication cycle: An in-depth morphological analysis. *Sci. Rep.* **8**, 4000 (2018).
42. Y. Liang *et al.*, Metagenomic analysis of the diversity of DNA viruses in the surface and deep sea of the South China sea. *Front. Microbiol.* **10**, 1951 (2019).
43. D. De Corte *et al.*, Viral communities in the global deep ocean conveyor belt assessed by targeted viromics. *Front. Microbiol.* **10**, 1801 (2019).
44. K. Subramaniam *et al.*, A new family of DNA viruses causing disease in Crustaceans from diverse aquatic biomes. *MBio* **11**, e02938-19 (2020).
45. S. W. Wilhelm *et al.*, A student's guide to giant viruses infecting small eukaryotes: From Acanthamoeba to Zooxanthellae. *Viruses* **9**, 46 (2017).
46. N. Chelkha *et al.*, A phylogenomic study of *Acanthamoeba polyphaga* draft genome sequences suggests genetic exchanges with giant viruses. *Front. Microbiol.* **9**, 2098 (2018).
47. F. Maumus, G. Blanc, Study of gene trafficking between Acanthamoeba and giant viruses suggests an undiscovered family of amoeba-infecting viruses. *Genome Biol. Evol.* **8**, 3351–3363 (2016).
48. M. Krupovic, D. H. Bamford, Double-stranded DNA viruses: 20 families and only five different architectural principles for virion assembly. *Curr. Opin. Virol.* **1**, 118–124 (2011).
49. E. S. Reynolds, The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J. Cell Biol.* **17**, 208–212 (1963).
50. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
51. J. Schindelin *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
52. J. Besemer, A. Lomsadze, M. Borodovsky, GeneMarkS: A self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29**, 2607–2618 (2001).
53. L. H. Bajrai *et al.*, Isolation of yasminevirus, the first member of klosneuvirinae isolated in coculture with vermamoeba vermiformis, demonstrates an extended arsenal of translational apparatus components. *J. Virol.* **94**, e01534-19 (2019).
54. D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).
55. D. G. Reteno *et al.*, Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* **89**, 6585–6594 (2015).
56. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
57. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
58. K. Katoh, D. M. Standley, A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**, 1933–1942 (2016).
59. P. Tommaso *et al.*, T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* **39**, W13–W17 (2011).
60. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
61. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
62. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
63. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
64. J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638 (2016).

Archives of Virology (2022) 167:711–715
https://doi.org/10.1007/s00705-021-05326-1

VIROLOGY DIVISION NEWS

"*Yaraviridae*": a proposed new family of viruses infecting *Acanthamoeba castellanii*

Paulo Victor de Miranda Boratto[1] · Graziele Pereira Oliveira[1] · Jônatas Santos Abrahão[1]

Published online: 9 January 2022

Abstract

Here, we propose the creation of the family "*Yaraviridae*", a new taxon to classify a virus infecting *Acanthamoeba castellanii* cells. Recently, we described the discovery of a new virus infecting free-living amoebae, yaravirus, which has features that strongly differ from those of all other viruses of amoebae described to date. Yaravirus particles are about 80 nm in diameter and have a dsDNA genome of ~45 kbp containing 74 ORFs, most of which (>90%) have no homologs in current databases. Together, these data support the creation of a new species ("*Yaravirus brasiliense*"), a new viral genus (here proposed as "*Yaravirus*"), and a new viral family (here proposed as "*Yaraviridae*") to classify yaravirus and other related viruses that may be described in the future. All of them are to be included into the existing realm *Varidnaviria* and the kingdom *Bamfordvirae,* due to the presence of a major capsid protein containing a double jelly-roll fold.

Introduction

Dozens of viruses infecting free-living amoebae have been described in different environments and parts of the world in recent years [1–6]. In addition to several other characteristics, these viruses are predominantly distinguished by the large size of their particles and a genome encoding a large number of proteins, a large proportion of which are ORFans, proteins that do not have any homologues in current protein sequence databases, [3, 6]. Recently, we described the discovery of a new virus infecting free-living amoebae, yaravirus, which has features that strongly differ from those of all other viruses infecting free-living amoebae described to date [6]. Yaravirus is the first viral isolate infecting *Acanthamoeba castellanii* cultures that potentially does not belong to the nucleocytoplasmic large DNA viruses (NCLDVs). This virus harbors a dsDNA genome with an exceptionally large number of ORFans (more than 90%), including novel proteins related to the construction of capsids [6]. Due to the lack of phylogenetic information, its origins are still a mystery, and similarity analysis with other viruses is unfeasible.

This virus was isolated from samples of muddy water collected from a creek near Lake Pampulha and named after an important character from the mythological stories of the Tupi indigenous tribes (Yara, the mother of waters). The novel characteristics of this virus led us to prepare and submit a formal taxonomic proposal for creation of a new species "*Yaravirus brasiliense*", a new viral genus (here proposed as "*Yaravirus*"), and a new viral family (here proposed as "*Yaraviridae*") to classify yaravirus.

Morphological properties

Yaravirus has a icosahedral capsid measuring about 80 nm in diameter (Fig. 1A). Cryo-electron microscopy images of purified particles suggest that yaravirus particles contain two capsid shells consisting of 26 viral proteins, one of which potentially represents a divergent capsid protein with a predicted double jelly-roll domain. This suggests that yaravirus potentially represent the first isolate of a novel group of amoebal viruses isolated from *Acanthamoeba castellanii* cells.

Handling Editor: Sead Sabanadzovic.

✉ Jônatas Santos Abrahão
jonatas.abrahao@gmail.com

[1] Laboratório de Vírus, Instituto de Ciências Biológicas, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil

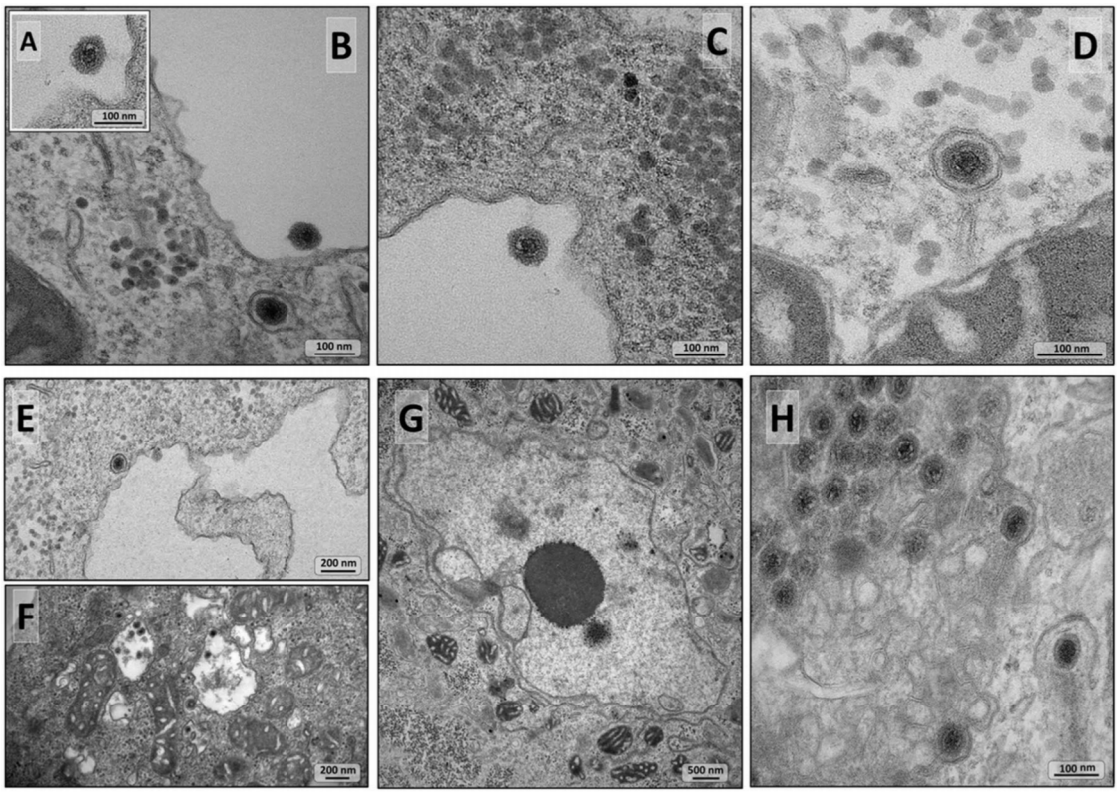


Fig. 1 **A** Transmission electron micrograph of a yaravirus virion. **B** Yaravirus particles attached to the outside part of the host cell membrane. **C** Yaravirus particles undergoing endocytosis. **D** Yaravirus particles incorporated by the amoeba inside an endocytic vesicle. **E** Yaravirus particles occurring individually in the interior of endocytic vesicles. **F** Yaravirus particles grouped inside endocytic vesicles. **G** Cytoplasm of the host cell showing mitochondrial recruitment while the viral factory develops. **H** Mature yaravirus particles being packaged inside vesicle-like structures at a late stage of infection, suggesting exocytosis as a mechanism for the release of the virions

Prevalence and host range

The distribution and natural hosts of yaravirus remain uncertain. This virus was isolated from samples of muddy water collected from a creek near Lake Pampulha, an artificial lagoon in the city of Belo Horizonte, Brazil (19°51 0.04S and 43°58 46.00W), in June 2015. The virus was isolated in amoebas of *Acanthamoeba castellanii* strain Neff (ATCC 30010). After isolation, experiments were performed by infecting *Acanthamoeba castellanii* cells with yaravirus at a low multiplicity of infection (MOI), given the fastidious nature of this virus. Yaravirus is not able to infect and establish a productive cycle *in vitro* in *Acanthamoeba polyphaga* or *Vermamoeba vermiformis*.

Properties in culture

Infection with yaravirus causes a decrease in the number of viable host cells while the viral genome is replicating. During the first hours of infection, *A. castellanii* cultures continue to grow until about 24 hours postinfection (hpi). Replication of the viral genome then progresses slowly, and cell lysis occurs only after 72 hpi Between 96 hpi and 7 days postinfection, there is no change in the amount of yaravirus genomic DNA detected, and lysis seems to stop, with the remaining trophozoites turning into cysts. In contrast to what has been observed with other viruses infecting free-living amoebae, these results suggest participation of a host receptor in the internalization of viral particles, since, at the beginning of infection, yaravirus particles are found to be attached to the outside part of

the amoebal plasma membrane (Fig. 1B–C). This is then followed by penetration of virus particles into the cell. Virions may be internalized individually or grouped inside endocytic vesicles, which, at a later stage of infection, are found next to a region previously occupied by the nucleus (Fig. 1D–F). After an eclipse phase, small crescents start to appear in the electron-lucent region of the virus factory, which is surrounded at this stage by several recruited mitochondria (Fig. 1G). Particles then develop progressively as they migrate toward the periphery of the cell. Sometimes, several yaravirus particles are packed into vesicle-like structures, suggesting that they are released by exocytosis, as is observed with other amoebal viruses (Fig. 1H). The remaining virions are released by lysis of the host cell. Although yaravirus is able to replicate in *Acanthamoeba*, considering its fastidious growth in this amoeba, it is possible that a different organism might act as its natural host.

Genomic and proteomic features

Yaravirus has a linear, double-stranded DNA genome, with a length of approximately 45 kbp, encoding 74 proteins. Its GC content is one of the highest among the known amoebal viruses, around 57%. Interestingly, as opposed to what has been observed for some members of the NCDLV group, the yaravirus genome does not seem to be enriched for any sequence motifs that might indicate the presence of a conserved promoter. In an intergenic region located between genes 29 and 30, there is a genomic island composed of six tRNAs (tRNA-Ser [gct], tRNA-Ser [tga], tRNA-Cys [gca], tRNA-Asn [gtt], tRNA-His [gtg], and tRNA-Ile [aat]). This is quite a remarkable number, given that a similar number of tRNAs is found in amoebal viruses with genomes much larger than that of yaravirus.

The most striking feature, however, is the number of genes with no similarity to sequences already described in databases. Genomic analysis at the nucleotide level revealed that there were no currently recognizable sequences in the yaravirus genome. The sequence similarity observed at the amino acid level is similarly low, with only six genes with distant matches (about 25–44% identity) in the nr databases. When adopting the same criteria used in genomic analysis of other amoebal viruses, 90% of the genome of the yaravirus is found to consist of sequences never described before by science. In an alternative strategy involving genome annotation based on analysis of the predicted three-dimensional structures of yaravirus proteins, the fraction of genes defined as ORFans was reduced to 80%. However, the proportion of unrecognizable proteins is still very impressive, reinforcing the idea that yaravirus is the first member of a newly discovered viral group.

These results were corroborated by proteomic analysis. The data show that yaravirus virions consist of 26 proteins. Comparisons between their predicted three-dimensional structures and structures currently available in databases have established a potential function for only five of them. Again, as observed for the genome, the proteome of yaravirus is primarily composed of proteins that have not been described previously.

Interestingly, despite representing a potential member of a newly discovered viral group, the yaraviruses seem to be exceedingly rare in nature. No sequences from the yaravirus major capsid protein (MCP) were retrieved in a search of 8,535 publicly available metagenomes generated from samples from diverse habitats across our planet. For sequences covering the yaravirus ATPase (NCVOG0249), only distant homologues in a group of endemic viruses of *Pleurochrysis* sp. were found, with only 33.9% identity to sequences from the same metagenomic databases.

Phylogenomics

Due to the large number of ORFan genes in the yaravirus genome, most of its proteins lack the phylogenetic information necessary to incorporate this virus into an existing taxonomic group. However, as discussed above, some of its genes have sufficient similarity to genes in the current databases to be used for phylogenetic analysis. For some genes, such as gene 02 (nuclease/recombinase) and gene 69 (bifunctional DNA primase/polymerase), yaravirus groups with members of the domain Eukarya (Fig. 2A and B). For other genes, as observed for gene 46 (hypothetical protein) and gene 40 (virion packaging ATPase), yaravirus seems to be related to sequences found in giant viruses (Marseillevirus strains and members of the family *Mimiviridae*, respectively) (Fig. 2C and D]. A structure-guided phylogenetic analysis was also performed for hits of the yaravirus MCP protein among 235 reference genome sequences of members of the phylum *Nucleocytoviricota* (plus contigs of three endemic viruses of *Pleurochrysis*). The alignments showed an interesting, well-supported clade, grouping the yaravirus MCP with metagenomic sequences from the distant endemic viruses of *Pleurochrysis* sp. (Fig. 3). Further attempts were made to find other similarities between these viral groups, but only a small number of protein candidates were found, with low sequence similarity (around 24 to 33% identity). These results highlight the great differences observed between yaravirus and other amoebal viruses, supporting the creation of a new taxonomic group to include this virus.

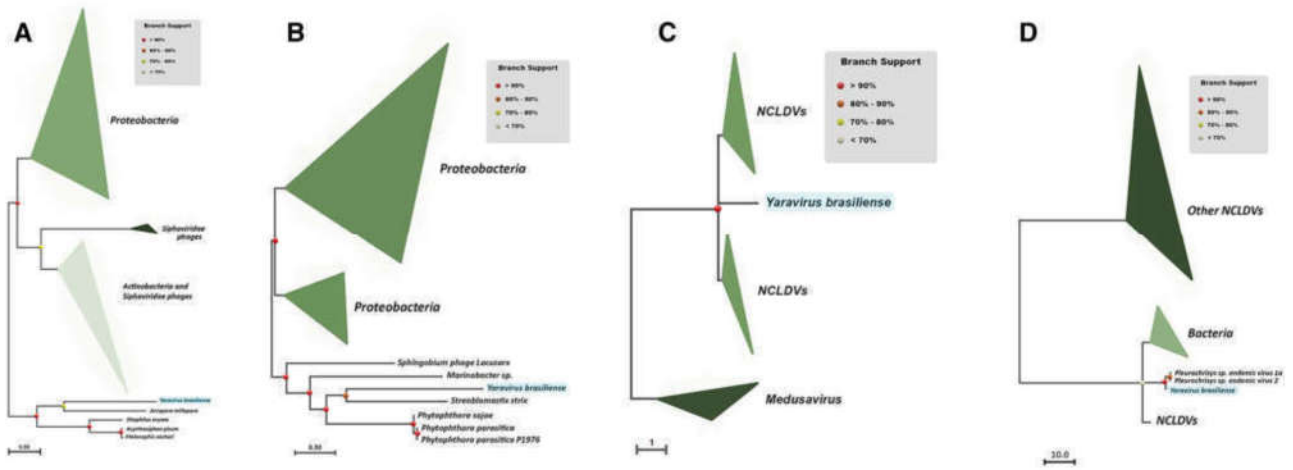


Fig. 2 Maximum-likelihood phylogenetic trees based on amino acid sequences of putative yaravirus proteins showing similarity to other amino acid sequences currently available in public databases. For each tree, it was constructed using amino acid sequences of yaravirus and the corresponding sequences belonging to bacteria, archaea, eukaryotes, and other members of the virosphere. Sequences were aligned using the tool ClustalW, available in MEGA X and performing 1,000 bootstrap replicates [7]. **A** Predicted exonuclease/recombinase. **B** Bifunctional DNA primase/polymerase. **C** A hypothetical protein. **D** Virion packaging ATPase. Yaravirus is highlighted in blue, and branch support is shown as coloured circles.

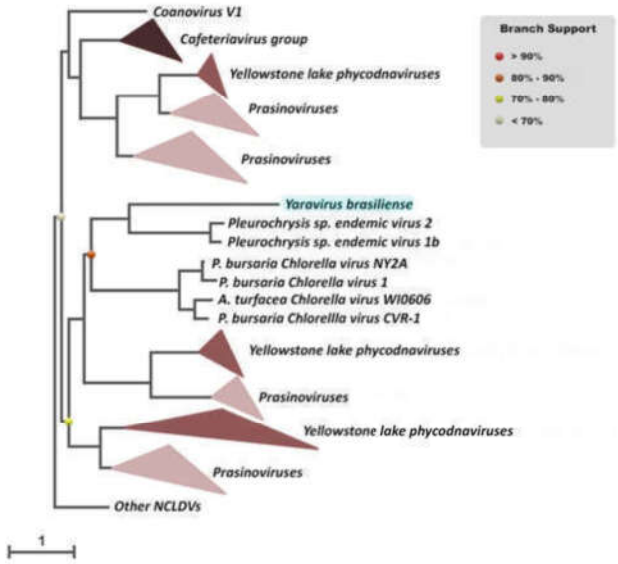


Fig. 3 A structure-guided phylogenetic analysis performed on hmm search hits for the putative yaravirus MCP protein among 235 reference genome sequences (from members of the phylum *Nucleocytoviricota* plus contigs of three endemic viruses of *Pleurochrysis*) and using hidden Markov models. The alignments were made using the tool Expresso, implemented the software T-Coffee using PDB structure files. Phylogenetic trees were constructed using the software IQ-Tree (v1.6.12; ref 7) with the LG+F+R8 model based on the built-in model select feature [8] and 1,000 ultrafast bootstrap replicates [9]. Yaravirus is highlighted in blue, and branch support is shown as coloured circles.

Conclusion

Yaravirus is a novel amoebal virus whose genome is mostly composed of ORFans. Although some of its genes show similarity to sequences already available in databases, the relationships are distant. In addition, yaravirus possesses other features that distinguish it from previously isolated viruses of amoebae. Its virion is about 80 nm in diameter, and its genome, with a size of ~45 kbp, contains 74 ORFs. Together, these data support the creation of a new species for which we propose the name "*Yaravirus brasiliense*", a new viral genus (here proposed as "*Yaravirus*"), and a new viral family (here proposed as "*Yaraviridae*"), in which yaravirus and other related viruses can be properly classified. The presence of a major capsid protein containing a double jelly-roll fold would warrant the inclusion of this virus in the existing realm *Varidnaviria* and the kingdom *Bamfordvirae*. A formal taxonomic proposal has been prepared and submitted to the ICTV and is currently under consideration and awaiting a ratification vote.

Acknowledgements We would like to thank Dr. Bernard La Scola (Aix Marseille Université) for the fruitful collaboration on topics related to amoebal viruses in the last years. We also thank colleagues from the Microscopy Center—UFMG and Laboratório de Vírus—UFMG for their technical support. We also thank FAPEMIG, CNPq, CAPES, and Pro-Reitoria de Pesquisa e Pós-Graduação-UFMG for financial support and scholarships. JSA is a CNPq researcher.

Funding We thank FAPEMIG (Grant No. PPM-00732-18), CNPq, CAPES, and Pro-Reitoria de Pesquisa e Pós-Graduação-UFMG for financial support and scholarships (Grant No. 02081/2018-6). JSA is a CNPq researcher.

Declarations

Conflict of interest The authors declare no competing interests.

Ethics approval SISGEN numbers: A3DAB3F, A580BBD, and ABF23CC. SISBIO numbers: 34293 and 33326

References

1. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie JM, Raoult D (2003) A giant virus in amoebae. Science 299(5615):2033. https://doi.org/10.1126/science.1081867
2. 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(51):21848–21853. https://doi.org/10.1073/pnas.0911354106
3. Philippe N, Legendre M, Doutre G, Coute 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(6143):281–286. https://doi.org/10.1126/science.1239181
4. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, Lescot M, Poirot O, Bertaux L, Bruley C, Coute Y, Rivkina E, Abergel C, Claverie JM (2014) Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci USA 111(11):4274–4279. https://doi.org/10.1073/pnas.1320670111
5. Abrahao 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(1):749. https://doi.org/10.1038/s41467-018-03168-1
6. Boratto PVM, Oliveira GP, Machado TB, Andrade A, Baudoin JP, Klose T, Schulz F, Azza S, Decloquement P, Chabriere E, Colson P, Levasseur A, La Scola B, Abrahao JS (2020) Yaravirus: a novel 80-nm virus infecting *Acanthamoeba castellanii*. Proc Natl Acad Sci USA 117(28):16579–16586. https://doi.org/10.1073/pnas.2001637117
7. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096
8. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285
9. 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

5. Artigo número 05:

5.1 *A Brief History of Giant Viruses' Studies in Brazilian Biomes*

(doi: https://doi.org/10.3390/v14020191)

No ano de 2021 completou-se uma década desde que os primeiros estudos envolvendo vírus gigantes se iniciasse na América do Sul, e mais especificamente no Brasil. Desde então muito do nosso conhecimento sobre esses micro-organismos foi se atualizando, muitas famílias virais foram sendo descritas nos mais diversos ambientes, e com isso, novos grupos de pesquisa brasileiros foram incentivados a prospectar diferentes vírus de amebas com características cada vez mais intrigantes. Esse artigo de revisão conta uma breve história sobre esses dez anos de investigação envolvendo os vírus gigantes de ameba em terras brasileiras. Para isso, compilamos aqui as principais contribuições de independentes grupos de pesquisa, buscando sintetizar todo o conhecimento acumulado sobre a diversidade e a excepcionalidade desses organismos encontrados em nosso país. Esta revisão contou com a contribuição de vários grupos de pesquisa que trabalharam com vírus gigantes nos últimos 10 anos no Brasil.

viruses

MDPI

Review

A Brief History of Giant Viruses' Studies in Brazilian Biomes

Paulo Victor M. Boratto [1], Mateus Sá M. Serafim [1], Amanda Stéphanie A. Witt [1], Ana Paula C. Crispim [1], Bruna Luiza de Azevedo [1], Gabriel Augusto P. de Souza [1], Isabella Luiza M. de Aquino [1], Talita B. Machado [1], Victória F. Queiroz [1], Rodrigo A. L. Rodrigues [1], Ivan Bergier [2], Juliana Reis Cortines [3], Savio Torres de Farias [4], Raíssa Nunes dos Santos [5], Fabrício Souza Campos [5], Ana Cláudia Franco [5] and Jônatas S. Abrahão [1,*]

[1] Laboratório de Vírus, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Minas Gerais, Brazil; pvboratto@gmail.com (P.V.M.B.); mateusmserafim@gmail.com (M.S.M.S.); asawitt1997@gmail.com (A.S.A.W.); anapbio2@gmail.com (A.P.C.C.); azvdobruna@gmail.com (B.L.d.A.); neogaps@gmail.com (G.A.P.d.S.); isabellaaquino92@gmail.com (I.L.M.d.A.); bastostalita04@gmail.com (T.B.M.); victoriafq18@gmail.com (V.F.Q.); rodriguesral07@gmail.com (R.A.L.R.)

[2] Embrapa Pantanal, Corumbá 79320-900, Mato Grosso do Sul, Brazil; bergiercpap@gmail.com

[3] Departamento de Virologia, Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-590, Rio de Janeiro, Brazil; cortines@micro.ufrj.br

[4] Laboratório de Genética Evolutiva Paulo Leminsk, Departamento de Biologia Molecular, Universidade Federal da Paraíba, João Pessoa 58050-085, Paraíba, Brazil; stfarias@yahoo.com.br

[5] Laboratório de Virologia, Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Porto Alegre 90.050-170, Rio Grande do Sul, Brazil; engraissanunes@gmail.com (R.N.d.S.); camposvet@gmail.com (F.S.C.); anafranco.ufrgs@gmail.com (A.C.F.)

* Correspondence: jonatas.abrahao@gmail.com

Abstract: Almost two decades after the isolation of the first amoebal giant viruses, indubitably the discovery of these entities has deeply affected the current scientific knowledge on the virosphere. Much has been uncovered since then: viruses can now acknowledge complex genomes and huge particle sizes, integrating remarkable evolutionary relationships that date as early as the emergence of life on the planet. This year, a decade has passed since the first studies on giant viruses in the Brazilian territory, and since then biomes of rare beauty and biodiversity (Amazon, Atlantic forest, Pantanal wetlands, Cerrado savannas) have been explored in the search for giant viruses. From those unique biomes, novel viral entities were found, revealing never before seen genomes and virion structures. To celebrate this, here we bring together the context, inspirations, and the major contributions of independent Brazilian research groups to summarize the accumulated knowledge about the diversity and the exceptionality of some of the giant viruses found in Brazil.

Keywords: amoebae viruses; Brazilian isolates; giant virus; NCLDV; virosphere; virus diversity

check for updates

Citation: Boratto, P.V.M.; Serafim, M.S.M.; Witt, A.S.A.; Crispim, A.P.C.; Azevedo, B.L.d.; Souza, G.A.P.d.; Aquino, I.L.M.d.; Machado, T.B.; Queiroz, V.F.; Rodrigues, R.A.L.; et al. A Brief History of Giant Viruses' Studies in Brazilian Biomes. *Viruses* **2022**, *14*, 191. https://doi.org/10.3390/v14020191

Academic Editor: K. Andrew White

Received: 18 November 2021
Accepted: 15 January 2022
Published: 19 January 2022

1. Introduction

The description and characterization of the first amoebal giant virus (GV) in 2003, Acanthamoeba polyphaga mimivirus (APMV), raised important questions regarding the limits of the virosphere. These first findings revealed viral particles of about 700 nm, non-filterable through 0.2 µm pore size filters [1]. Although not the first described nucleocytoplasmic large DNA virus (NCLDV) (phylum *Nucleocytoviricota*), which includes other families such as *Poxviridae* [2], the original discovered member of the family *Mimiviridae* motivated new interpretations of crucial features in an organism recognized as a virus, advancing both knowledge of, and perspectives on, the most abundant group of organisms on Earth [3].

Much of the subsequent work on GVs has been driven by curiosity and the possibility of isolating novel groups of amoebal viruses and finding intriguing new characteristics. For instance, in 2008 La Scola et al. had previously isolated a distinct strain of APMV, the

acanthamoeba castellanii mamavirus, together with the first ever described virophage, the Sputnik virus (SNV), both found in a water-cooling tower of a hospital in France [4]. Here, the interest in describing more GVs contributed by influencing the consolidation of "virophages" as new satellite-like viruses, which are dependent on the mimivirus factory for their replication by putatively hijacking some key features (e.g., the viral RNA polymerase) [5].

Similarly, in 2009, Boyer et al. reported the isolation of the marseillevirus, a novel GV for which analysis of its core genes suggested a previously uncharacterized family of NCLDV. In addition, by unveiling some of the main features of the genome's repertoire for this new GV, the authors have proposed amoebas as potential "melting pots" of microbial evolution, given the convenient intracellular environment for gene transfer among parasites, including complex genomes that could advent from different GVs' viral sources [6]. Some of the GVs' hosts (different amoeba genus, e.g., *Acanthamoeba*) are indeed considered ubiquitous, found in almost all latitudes [7,8], as well as in a wide-range of environments, including wastewater [9], terrestrial and (deep) marine water [7,10], thermal springs [11], permafrost [12], ventilation and air conditioning systems, and even in hospital settings [13,14]. Notwithstanding, GVs can be found in a large set of different native hosts or host-associated organisms, from other various species of amoebas [8,15] to filtering feeding organisms such as oysters [16]. Recently, metagenomics studies have also indicated that GVs are even more abundant in marine environments than prokaryotes, suggesting that these viruses may play a fundamental role in nature as biological control agents, regulating biogeochemical cycles, and potentially acting as evolutionary driving forces [17–19]. Ultimately, hijacking or utilizing cellular components and translational machinery may indicate a common origin, regarding information on life's evolution, and the presence of translation proteins may open new hypotheses about GVs' origin and phylogenetic relationships with other domains of life [20].

The broad-spectrum environmental profile of GVs made Brazil an interesting field to search and study these microorganisms, especially considering the wide range and diversity of environments and biological dispersion throughout biomes, settings and native habitats around the country, such as the Amazon forest and Cerrado savannas, as well as the Pantanal wetlands, including the soda lakes of Nhecolândia in the middle of the woods, which hide rich organic sediments [7,10,21–23]. These characteristics were reflected in the findings and discoveries of several GV isolates in the Brazilian territory, such as: (i) tupanvirus soda lake, isolated from the Brazilian Pantanal (Nhecolândia lakes) and tupanvirus deep ocean, isolated from sediment at 3000 m below the water surface line at 'Bacia de Campos', in Rio de Janeiro [10]; (ii) samba virus (SMBV), isolated from the Brazilian Amazon [22]; (iii) cedratvirus getuliensis, from sewage samples in Minas Gerais state [24]; (iv) niemeyer virus [25], faustovirus mariensis [26] and yaravirus [27], all of them isolated from the Pampulha urban lagoon; (v) a number of *Mimivirus* isolates [7,21]; and many others (Figure 1A,B), some of which we will further discuss in this review.

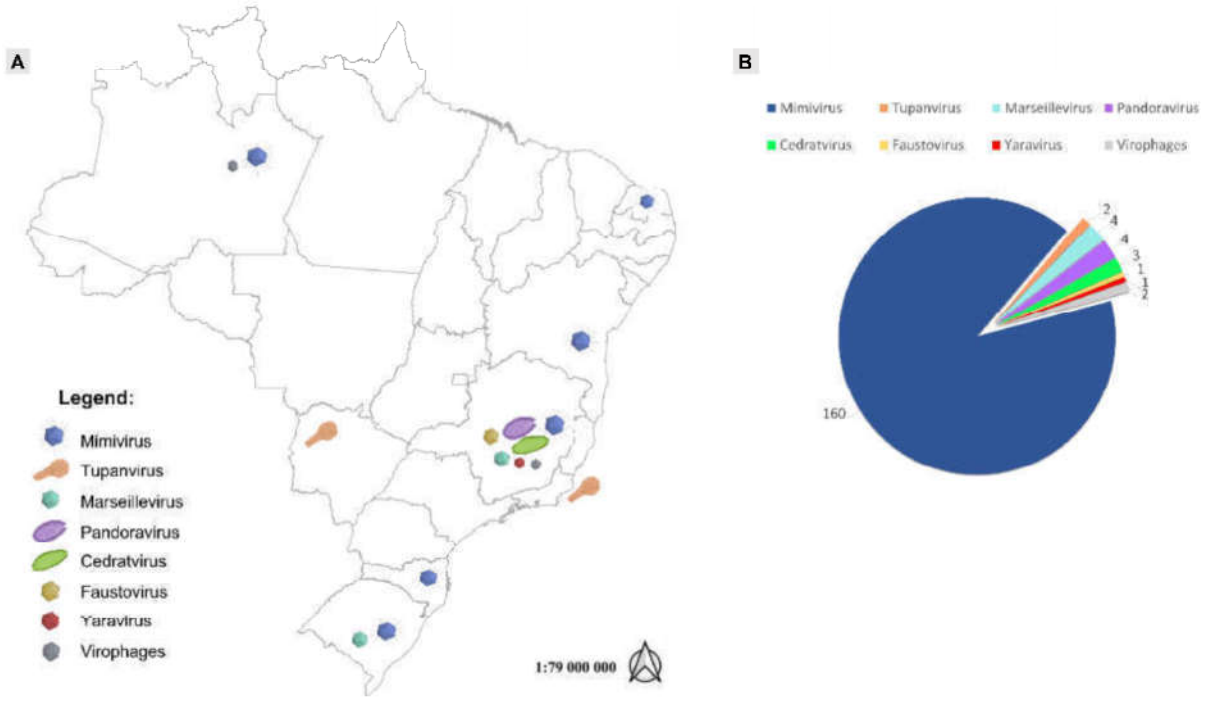


Figure 1. Location and numbers of giant viruses isolated in Brazil. (**A**) Schematic map showing the sites of isolation for the major groups of giant viruses discovered in Brazil. (**B**) Number of isolates discovered for each of these groups.

2. Giant Viruses Discovery and Isolation

2.1. Mimiviruses Boosted Amoebal Giant Viruses' Research

The first amoebal GV isolated in Brazil dates from 2011. During a field trip to the Brazilian Amazon, aiming to search for GVs, water samples were collected from the Negro River, in Manaus city. These samples were then assessed for prospection using *Acanthamoeba castellanii* cells as an isolation platform, which allowed the discovery of the first Brazilian GV, named samba virus (SMBV) [22].

Samba viruses have mimivirus-like particles, showing capsids with an average diameter of 527 nm, surrounded by fibrils of 155 nm [28]. The SMBV genome is composed of about 1.2 Mb, and the phylogenetic analysis clustered it within the lineage A of the mimiviruses (Table 1). Analysis of the SMBV replication cycle using a set of electron microscopy images showed several similarities with the APMV replication cycle. Moreover, these images also revealed the presence of smaller viral particles that were further confirmed as the first Brazilian virophage named Rio Negro virophage (RNV) [22]. A few years later, RNV genome was sequenced and assembled presenting 18,145 bp, very similar to the sputnik 2 virophage genome [29].

Later, a new sputnik-like virophage named guarani virophage was also isolated from water samples obtained in the Pampulha Lagoon, Belo Horizonte city, Minas Gerais state. A deep characterization of its genome (18,967 bp) was performed, and its replication is described as rather slow (replication at 4 h.p.i. and particles morphogenesis at 16 h.p.i.) when compared to the cycle of its associated GV [30].

After SMBV discovery, several mimiviruses belonging to the three currently known lineages (A, B and C) were isolated from different Brazilian environmental samples, such as the so-called "Br-mimiC", mimivirus golden (MVGD), isolated from golden mussels (*Limnoperna fortunei*) from Guaíba Lake, Rio Grande do Sul, in 2014 [31] and mimivirus gilmour (MVGM), isolated from water collected at the Pampulha Lagoon, in 2015 [21]. In this same work, the isolation of another 64 mimiviruses were described from water samples

collected at the Pampulha Lagoon. They were obtained from three different *Acanthamoeba* species (*A. castellanii*, *A. polyphaga* and *A. griffini*), and had representatives in the three lineages of mimiviruses [21]. Also in 2015, 20 mimiviruses belonging to the lineage A were obtained from oyster-related samples of three different coastal regions of Brazil [16]. Considering their water-filtering capacity, these bivalves were tagged as excellent sources for the isolation of new GVs because their body allows the accumulation of both viruses and amoebas [16].

Table 1. General features of Brazilian giant viruses with complete sequenced-genomes.

Group of Virus	Virus	Type of Sample	Location (Year of Isolation)	Genome Size (bp)	ORFs	ORFans	GC %	Reference
Mimiviridae (lineage A mimivirus)	Samba virus	Fresh water	Negro River (2011)	1,181,380	971	0	27	Campos et al., 2014
	Amazonia virus	Fresh water	Negro River (2011)	1,179,119	979	1 (0.1%)	27	Assis et al., 2015
	Kroon virus	Urban lake water	Lagoa Santa city (2012)	1,221,932	944	3 (0.3%)	27	Assis et al., 2015
	Oyster virus	Oysters	Santa Catarina state (2013)	1,200,220	948	1 (0.1%)	27	Assis et al., 2015
	Niemeyer virus	Urban lake water	Pampulha Lagoon (2011)	1,299,140	1003	0	28	Boratto et al., 2015
Mimiviridae (lineage B mimivirus)	Borely moumouvirus	Fresh Water	Serra do Cipó (2018)	1,038,187	947	3 (0.3%)	25.2	Silva et al., 2020
Mimiviridae (lineage C mimivirus)	Mimivirus gilmour	Urban lake water	Pampulha Lagoon (2014)	1,258,663	1135	28 (2.4%)	26	Assis et al., 2017
	Mimivirus golden	Golden mussels	Guaíba Lake (2014)	1,248,960	1127	19 (1.6%)	26	Assis et al., 2017
Mimiviridae	Tupanvirus deep ocean	Deep Ocean sediments	Campos dos Goytacazes city (2018)	1,439,508	1276	378 (29.6%)	28	Abrahão et al., 2018
	Tupanvirus soda lake	Soda Lake	Nhecolândia, Pantanal biome (2018)	1,516,267	1359	375 (27.6%)	28	Abrahão et al., 2018
Marseilleviridae	Brazilian marseillevirus	Sewage	Pampulha Lagoon (2014)	362,276	491	29 (5.9%)	43.3	Dornas et al., 2016
	Golden marseillevirus	Golden mussels	Guaíba Lake (2014)	360,610	483	43 (8.9%)	43.1	Santos et al., 2016
Cedratviruses	Brazilian cedratvirus	Water supplemented with biofloc	Belo Horizonte city (2018)	460,038	533	11 (2.1%)	42.9	Rodrigues et al., 2018
Faustovirus	Faustovirus mariensis	Urban lake water	Pampulha Lagoon (2019)	466,080	483	0	36	Borges et al., 2019
Yaravirus	Yaravirus brasiliensis	Muddy water	Pampulha Lagoon (2020)	44,924	74	68 (91.9%)	57.9	Boratto et al., 2020

In another study, a mimivirus-related isolate called Niemeyer virus (NYMV) was discovered, once again from water samples obtained from the Pampulha Lagoon [25]. NYMV has a genome of approximately 1,299,140 bp, harboring a set of duplicated aminoacyl-tRNA synthetases, which suggests that such duplications may be important for the evolutionary history of mimiviruses (Table 1). In 2017, another lineage A mimivirus was described, this time from water samples collected from an urban lake at the Lagoa Santa city, also in the Minas Gerais state, and it was named kroon virus (KV) (Figure 2A) [32]. The study of KV (1,221,932 bp genome [33]) has established an interesting view of the distinct ways by

which the major capsid protein (MCP) mRNA can be differentially processed, depending on the lineages of mimiviruses (Table 1) [32]. Apparently, for each of these mimiviruses' lineages there is a genetic layout concerning how the MCP gene is organized in terms of its exons/introns and how they are arranged. As an example, in the KV study the nucleotide sequences of the third exon of the MCP (observed in the genome of all mimi-viruses) was described to be an alternative marker to disentangle each of the three lineages. In addition, a different form of mature mRNA was also described in transcripts of the MCP for mimiviruses of a given lineage (e.g., APMV and KV) [32]. Subsequently, in 2018, 64 giant viruses of the *Mimiviridae* family (26 from lineage A, 13 from lineage B, two from lineage C and 23 from unidentified lineages) were described from various types of samples, including marine water from Antarctica, which was the first time to our knowledge that mimiviruses were isolated in this continent [7].

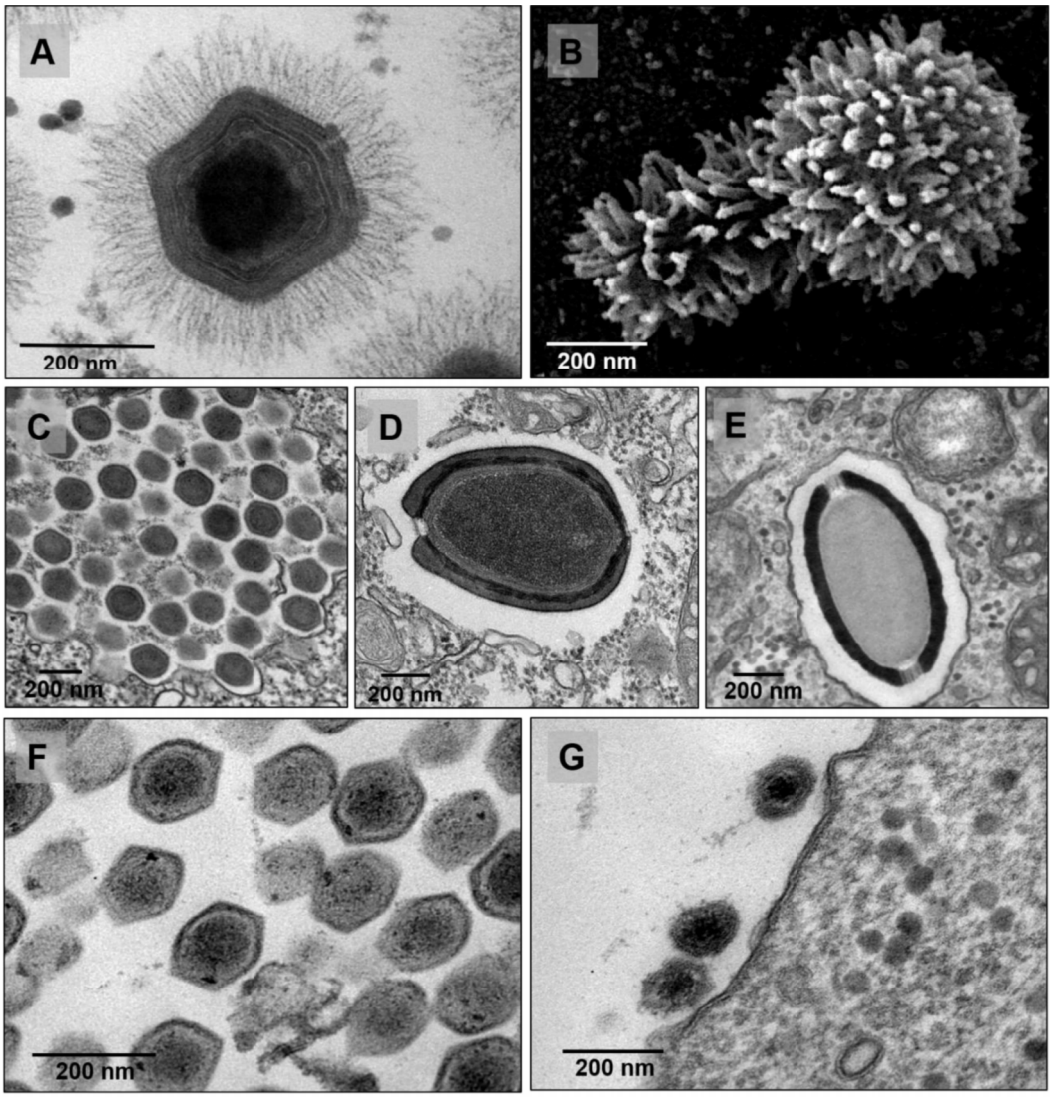


Figure 2. Panel with TEM images for the major groups of amoebal viruses isolated in Brazil. (**A**) mimivirus, (**B**) tupanvirus (source: 10.1038/s41598-018-36552-4), (**C**) marseillevirus, (**D**) pandoravirus, (**E**) cedratvirus, (**F**) faustovirus, and (**G**) yaravirus.

The year 2018 has also been marked by the description of one of the longest and most complex viruses described to date, obtained from a set of samples collected from

extreme aquatic environments. The tupanvirus soda lake was isolated from samples collected from an alkaline salty lake (Nhecolândia, Pantanal, Brazil) while the deep ocean tupanvirus was obtained from 3000 m depth sediment samples below the Atlantic Ocean at 'Campos dos Goytacazes' (Figure 2B) [10]. In contrast to other giant viruses previously isolated, these GVs are able to infect and establish a productive cycle in many species of amoebas, including *Acanthamoeba* spp., *Vermamoeba vermiformis*, *Dictyostelium discoideum* and *Willartia magna* [10]. The tupanviruses' capsid is similar to that of other mimiviruses already described in this review, around 450 nm in diameter, with a pseudo-icosahedral symmetry, covered by a layer of fibrils. They also present a "stargate" vertex, that is, a noticeable star-shaped opening at one capsid vertex. Interestingly, however, these viruses have a tail attached to the capsid, which is also covered with fibrils, a feature never seen before for amoebal viruses (Figure 2B). Due to the plasticity of this tail, the particles can vary from 1.2–2.3 µm in length, making it the longest virus ever described [10].

The genomes of the tupanviruses are complex and composed of double-stranded DNA of 1.44–1.51 Mb, encoding 1276–1425 predicted proteins (Table 1) [10]. Phylogenetic analysis using the DNA polymerase B family gene and other unique features exhibited by these viruses suggested that the tupanviruses group together with other mimiviruses form a distinct clade, which supported the proposal to form a new genus called "*Tupanvirus*" [10,34]. These viruses have been shown to be even more surprising, as deep genome analysis detected the largest translational apparatus ever described in the virosphere, with 20 aminoacyl-tRNA synthetases (aaRS), 67–70 tRNAs, in addition to other proteins in the translation process, such as translation factors (initiation, elongation and release) and proteins related to tRNA and mRNA maturation [10,34,35]. In addition, 20% of their genome is similar to genes originating from cellular organisms, with 9% from eukaryotes (of these, 3% originate from amoebas), 3% from archaea and 8% from bacteria [35].

These findings support data that demonstrate how other groups of organisms are relevant in studying the evolution of NCLDV genomes (Table 1). The fact that they have these genes shared with members of other cellular domains suggests that tupanviruses could also be found in non-extreme environments [35]. Altogether, the genetic arsenal of these and other mimiviruses within the virosphere add new levels of complexity to the understanding of the tree (or rhizome [36]) of life [20,37].

2.2. The Second Family Arises: The Discovery of Marseilleviruses

After the discovery of the first mimiviruses, the search for GVs intensified. In 2009, a virus named *Marseillevirus marseillevirus* was isolated in a biofilm from a water cooling tower in Paris, France [6], which gave rise to the family *Marseilleviridae*, officially recognized by the International Committee on Taxonomy of Viruses (ICTV) in 2013 (Figure 2C) [38]. Since then, other marseilleviruses have been isolated from different sources: (i) Lausannevirus, was discovered in water samples collected from the Seine river, in 2011 [39]; (ii) Cannes 8 virus was isolated from water in a cooling tower in Cannes, in 2013 [40]; (iii) tunisvirus and Fontaine Saint-Charles virus were isolated from freshwater collected in decorative fountains in Ariana, Tunisia, and in France, respectively [41,42]; (iv) insectomime virus was isolated from the internal organs and digestive tract of a dipteran drone fly's larvae [43]; (v) Senegalvirus was discovered during metagenomic analysis of the bacterial diversity in the human gut microbiota from a apparently healthy African individual, in 2012 [44,45]; (vi) In 2014, the genomic characterization of Melbournevirus was reported, isolated from a freshwater pond in Melbourne, Australia [46]; and (vii) Port-Miou virus, isolated from a sample from a brackish submarine spring, in the Cassis Port-Miou Calanque, France, in 2015 [47].

Furthermore, different phylogenetic lineages of marseillevirus have been described. Initially, the phylogenetic analysis suggested the existence of three distinct lineages: Lineage A, consisting of *Marseillevirus*, Cannes 8 virus, Senegalvirus and Melbournevirus; Lineage B, consisting solely of Lausannevirus; and Lineage C, consisting of tunisvirus and insectomime virus [42]. That was based on phylogenetic reconstructions carried out with core genes

including the NA polymerase B family, the VV A18 helicase, the D5 primase–helicase, the very late transcription factor 2B and the MCP [42].

However, the discovery of the first marseillevirus in America resulted in the creation of a new lineage in the family. The Brazilian Marseillevirus (BrMV) was described in 2014 from a sewage sample from a treatment station in the Pampulha lagoon [47]. The new lineage is supported by comparative genomic analyses highlighting several divergences between BrMV and other marseilleviruses (Figure 3) [47].

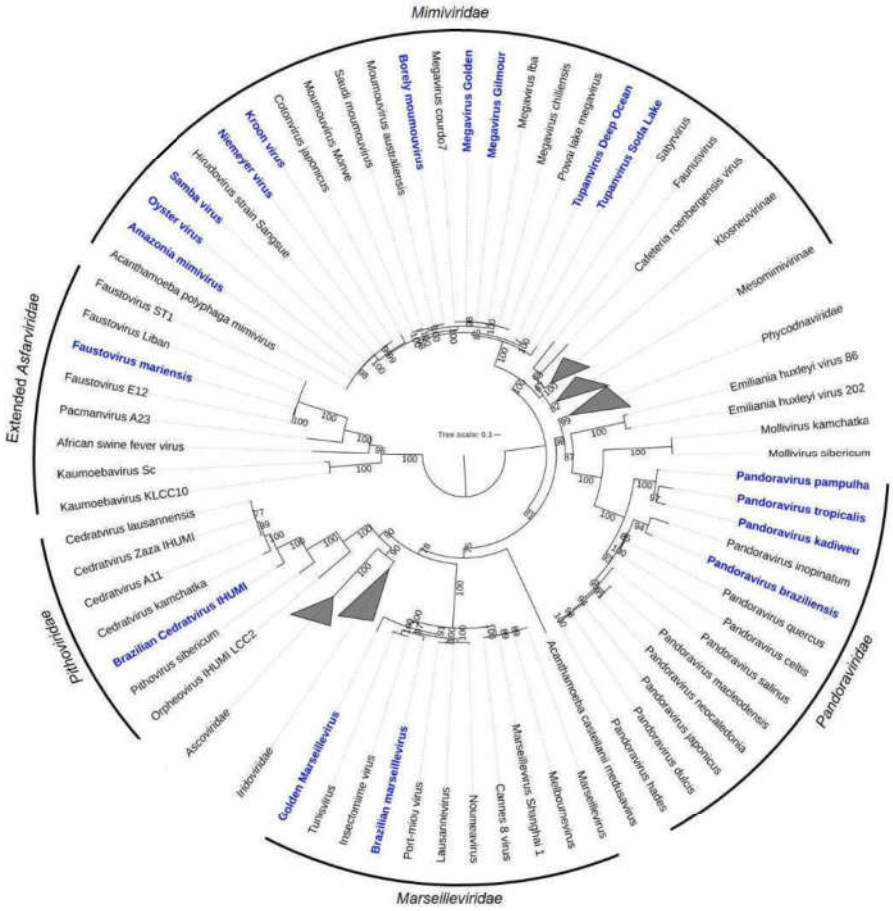


Figure 3. Maximum likelihood phylogenetic tree based on amino acid sequences of DNA polymerase B family of *Nucleocytoviricota*. Brazilian isolates are bold and highlighted in blue. Sequences were aligned using Muscle [48] and low conserved regions were removed using trimAl [49]. The tree was built using IQ-TREE [50] with 1000 ultrafast bootstrap replicates and the VT+F+R7 model chosen by ModelTest according to Bayesian Information Criterion. The tree was visualized in iToL [51]. The tree scale indicates the substitution rate.

A few years later, in 2016, the golden marseillevirus (GMar) was described as a new member isolated from golden mussels collected in southern Brazil [52]. The structure of the virus particles strongly resembled other marseilleviruses, with particles of approximately 200 nm, obtained from a coculture with A. polyphaga. The genome is composed of a circular dsDNA with 360,610 bp, comparable in size to the genomes of other members of the family *Marseilleviridae*, which range from 346,754 bp to 386,631 bp for lausannevirus and insectomime virus, respectively (Table 1) [52]. A total of 483 ORFs were characterized. Curiously, despite this genome size similarity, the GMar genes' content harbors 48.03% uncharacterized proteins. Many of these uncharacterized proteins can be considered as orphan genes (ORFans), also reported in other GVs such as Pandoravirus with

93% ORFans [53]. In addition, comparatively to the 212 genes shared among Brazilian marseillevirus, marseillevirus, lausannevirus, tunisvirus and golden marseillevirus, there are fourteen non-shared genes, of which seven are among the GMar genes [52].

2.3. *Opening the GVs' Box: The Discovery of Pandoraviruses*

Ten years after the discovery of the first GVs, the description of a brand-new group of amoebal viruses has led virologists to become again surprised, as a series of new paradigms started to be challenged and the study of modern virology advanced. At the time, the discovery and characterization of the pandoraviruses established for the first-time a group with viral particles with sizes as great as 1 µm in length and genomes that exceeded the mark of 2.5 Mb, with an astonishing number of 93% of genes without recognizable homologs in available databases (e.g., GenBank) (Figure 2D) [53].

Before the investigations in Brazil, it is important to mention the initial studies regarding the first representatives of this group. Starting from 2013, these discoveries were made: (i) *Pandoravirus salinus*, isolated from a superficial sediment layer collected at the mouth of the Tunquen river in Chile; and (ii) *Pandoravirus dulcis*, isolated from a mud taken at the bottom of a freshwater pond near Melbourne, Australia [53]. A couple of years later, a study led to the reinvestigation of an endosymbiont isolated from an *Acanthamoeba* strain and concluded, by whole genome sequencing, that this organism was in fact a pandoravirus isolate, named Pandoravirus inopinatum [54,55]. In another work, a newly characterized isolate called Pandoravirus celtis was used to investigate a putative scenario in which the genetic divergence among the different isolates of pandoraviruses was caused by an ability of these viruses to perform the creation of genes through a de novo microevolution process [56].

During the years 2018 to 2019, two different studies carried out a series of in-depth approaches focusing on establishing a detailed view of the diversity of pandoraviruses, their evolution processes and aspects of their replication cycle [57,58]. In the first study, three samples of pandoravirus were first isolated and named as pandoravirus quercus, pandoravirus neocaledonia and pandoravirus macleodensis. Their replication cycles were independently investigated and interestingly, for the first time, the mature particles of pandoraviruses were filmed while being exocytosed by vesicles which were full of viruses [57]. The genomes of these isolates were fully sequenced and a new stringent reannotation protocol was established. With this new methodology, the genetic analysis of different isolates suggested a still open pan-genome for GVs, in which each novel isolate is predicted to be responsible for contributing more than 50 additional genes [57].

For the second study, three novel Brazilian isolates were used: (i) pandoravirus kadiweu, coming from samples of water collected in the city of Bonito, Mato Grosso do Sul; (ii) pandoravirus pampulha, and (iii) pandoravirus tropicalis, both coming from samples of water from an artificial lake located at the city of Belo Horizonte, Minas Gerais [58]. Here, the microscopy analysis was an important tool, not only to reinforce some already established data but also to reveal new features of the virus replication. As for other GVs, within 30 min of infection the pandoravirus virions were phagocytosed and engulfed inside a host vesicle called the phagosome [57]. This structure quickly fuses with lysosome-like organelles and triggers the next stage of replication, which is the start of viral uncoating [53,57]. The next step involves an intense manipulation of the host cell and deep modification of the cytoplasm environment in order to make the region of viral morphogenesis, known as the viral factory. The loss of the cell nucleus and an intense recruitment of the host membranes and mitochondria are necessary for this. The beginning of viral morphogenesis does not seem to have a polarization, as thought earlier [53]. Finally, the viral cycle ends with the host cell lysis [53,57].

Interestingly, however, it was observed in some microscopy images that several pandoravirus particles were packaged inside vesicles and transported to the periphery of the host cell before amoebal disintegration. Additionally, one-step-growth curves have shown the beginning of viral release around 6 to 9 h post-infection, before the onset of

the amoebas' lysis [58]. These results, together with data that show a negative impact on pandoravirus release by cells treated with brefeldin (a membrane traffic inhibitor), suggest an important role of exocytosis for early liberation of pandoravirus particles in an amoeba infection [58]. Such observations are commonplace to other GVs with analogous replication cycles, including, for example, the cedratviruses described in the next section.

2.4. A Double-Corked GV: Isolation and Characterization of the Cedratviruses

Viruses belonging to the cedratvirus group were first detected in 2016, with the isolation of Cedratvirus A11, a viral representative coming from diverse environmental samples collected in Algeria [59]. Their structure is constructed by a ~1 µm ovoid-shaped particle, resembling some morphological features of the pithovirus virions, though with a notable difference: the presence of two corked regions (instead of a single one) at the extremities of the particle [59]. Their genome is composed of a circular dsDNA with about 590 kbp, and it has been found to share a close relationship with the genomes of the two currently known pithoviruses (both in size and in genome content), pithovirus sibericum and pithovirus massiliensis [59].

The second cedratvirus isolate, called cedratvirus lausannensis, was obtained in an attempt to look for amoebal-resisting bacteria inside a drinking water plant located at the Morsang-sur-Seine commune, in France [60]. Four other isolates have been discovered since then: (i) cedratvirus zaza IHUMI, deriving from samples of sterile distilled water collected near Toulon city, in France; (ii) Brazilian cedratvirus IHUMI, collected from water samples supplemented with bio-floc in Belo Horizonte city, Brazil; (iii) cedratvirus Kamchatka, obtained from a muddy grit soil collected next to a volcano area in Russia; and (iv) cedratvirus getuliensis (Figure 2E), collected from sewage samples from the Itaúna city, Brazil [24,61,62]. Interestingly, the isolate Brazilian cedratvirus IHUMI is a representative of the group which harbors both particle and genome sizes with remarkable differences in comparison with the other cedratviruses discovered so far. The virion is approximately 910 nm in length, with some of the particles reaching around 696 nm, and the genome is also smaller, with a DNA molecule of 460,038 bp (Table 1) [63]. Comparative genomic analysis also indicated that this Brazilian isolate is the founding member of a new lineage of cedratviruses (Figure 3) [63].

In 2018, through the analysis of a series of electron microscopy images and by performing biological assays, an interesting study has helped to reveal most of the steps in the replication of cedratviruses. As expected for an amoebal virus with large-sized virions, the viral cycle starts by the particles getting into the infected cell through the exploration of a phagocytic pathway that is physiologically presented by the host [24]. Corroborating this observation, lower titers of the cedratvirus virions are observed when the infected amoebal cells are pre-treated with cytochalasin D, an inhibitor of phagocytosis. The cycle then progresses to the formation of an electron-lucent viral factory (as large as the cellular nucleus) in the cytoplasm of the infected amoeba and, differently from observed during infection of most giant viruses, the cellular nucleus seems to remain intact [24].

However, some typical cellular alterations are still observed, such as the recruitment of mitochondria around the viral factory region, the polarization of lysosomal vesicles in the infected cell and an intense traffic of membranes which were seen to be important during the morphogenesis of cedratvirus virions [24]. This step is described as very complex and relies on the formation of several membrane precursors (or crescents) which later assume the correct conformation of a mature viral particle. Finally, the viral cycle ends with the mature particles released via cell lysis or exocytosis [24]. Cedratviruses also present structural similarities and infection features to other GVs, such as the orpheoviruses, as discussed below.

2.5. Another Amoeba, Another Virus: Discovery and Characterization of Orpheovirus

By implementing amoebas of the *Vermamoeba vermiformis* species as a platform of isolation, new groups of viruses were discovered from different samples. Among them, an

Orpheovirus was isolated in Marseille, France, from samples of rat stool [62]. Nevertheless, Souza et al. observed that, differently to previous findings of viruses infecting amoeba, CPE caused by Orpheovirus could be split into an early stage (3 to 12 h.p.i.), when cells stretch into a branched fusiform shape, and a late stage (starting at 24 h.p.i.), when cells become rounded [64].

The in-depth characterization of the replication cycle demonstrates that it takes approximately 30 h to be completed. It is suggested that one or more particles of Orpheovirus, which are around 1.1 µm, are phagocytized by the host cell within 1 h.p.i. [62,64]. After entry, the particle's internal content is released when the membrane that surrounds the viral core fuses with the endosomal membrane through a structure called ostiole, located at the apex of the particle. Subsequently, the formation of the large electron lucent viral factory is observed, concomitantly with the recruitment of mitochondria and membranes [64]. Membrane recruitment and bleb formation also seems to be important for the viral factory formation and particle morphogenesis since they are affected by treatment with a membrane trafficking inhibitor at the middle stage of infection (8 h.p.i.), which is also observed for cedratviruses [24]. Similarly, as described for other viruses, the particle morphogenesis initiates with the formation of electron-dense semicircular structures, which are filled with their internal content until the formation of the complete closed particle [64].

The complete particle presents smaller fibrils, when compared to mimiviruses, and at least two layers between the fibril layer and the inner membrane [62,64]. Finally, the infectious particles start to be released by exocytosis, detected in the supernatant at 12 h.p.i. Moreover, it is observed that cell lysis also plays a role in viral particle release, mostly at late timepoints of infection. Along with the infectious particles, the formation of defective particles is also observed [64].

2.6. *The Isolation and Characterization of Faustoviruses*

The faustoviruses are a group of giant viruses first detected in 2015 from samples of sewage from different regions in France and in Dakar, Senegal [65]. In Brazil, the first representative of this group was isolated and described in 2019, from prospecting studies of water samples from the Pampulha lagoon. Faustovirus mariensis, as it was called, is a virus with icosahedral particles reaching approximately 190 nm in diameter and inducing cytopathic effects on amoebas of the *Vermamoeba vermiformis* species (Figure 2F) [26]. Their genome is composed of a circular, double-stranded DNA molecule of about 466,080 bp (Table 1). Like other GVs, the f. mariensis replication cycle starts with the infection of the amoeba in its trophozoite form. This infection progresses to the formation of a large electron-lucent viral factory and the recruitment of mitochondria to its periphery [26]. The morphogenesis of f. mariensis is similar to that of other faustoviruses previously described in the literature, with new mature particles being formed in small honeycomb structures within the cytoplasm of the host cell. Lysis of the infected cell is the most important means of releasing the f. mariensis progeny described so far [26].

In a rare antiviral strategy described for GVs and their amoebal hosts, Borges et al. have observed that the infection of *Vermamoeba vermiformis* cultures is able to trigger a process of encystation of the neighboring cells, trapping the particles of f. mariensis inside their host and preventing further infection in the population of amoebas [26]. This event, considered to be observed for the first time in these viruses, was directly influenced by f. mariensis infection at a multiplicity of infection (MOI) dependent rates. When cysts were derived from cells infected at high MOIs, they were permanently incapable of excysting, therefore becoming trapped inside the particles of f. mariensis. However, when these amoebal cells came from infections at lower MOIs, only the cells with neither viral particles nor factories were able of excysting [26].

Faustoviruses are also phylogenetically related to kaumoebaviruses and asfarviruses, with the hypothesis that a common ancestor is shared between these viruses (Figure 3) [62]. After analysis considering this evolutionary proximity, motifs that play the role of promoter sequence in asfarvirus have been identified within the faustovirus genome, leading to the

conclusion that rich A-T (TATTT and TATATA) regions may also have an important role in the gene expression of both kaumoebavirus and faustovirus. These findings shed new light for a better understanding of giant virus's gene expression [66]. As aforementioned, intriguing information regarding the GVs' discovery and characteristics are quite common, and some unique factors have attracted attention in the field, such as the recent discovery of the yaravirus in 2020 [27].

2.7. *Yaravirus, a Small Virus among the Giants*

In late 2020, the discovery of a new lineage of dsDNA virus would enhance our knowledge on the diversity and evolution of viruses in amoeba. The yaravirus brasiliensis, as it was called, has been described as a novel virus of *Acanthamoeba castellanii*, harboring a genome of ~45 kbp enclosed in an icosahedral particle of about 80 nm in diameter [27]. Differently from any other virus isolated from acanthamoeba so far, this virus does not seem to share many of the features which are thought to represent the NCLDV, as it has neither a large particle nor a complex genome (Figure 4) [27]. This may indicate one of the following: (i) yaraviruses either belong to an extremely reduced group of amoebal viruses which are part of the NCLDVs; or (ii) these viruses represent the first discovered lineage of amoebal viruses that are not part of this complex group [27].

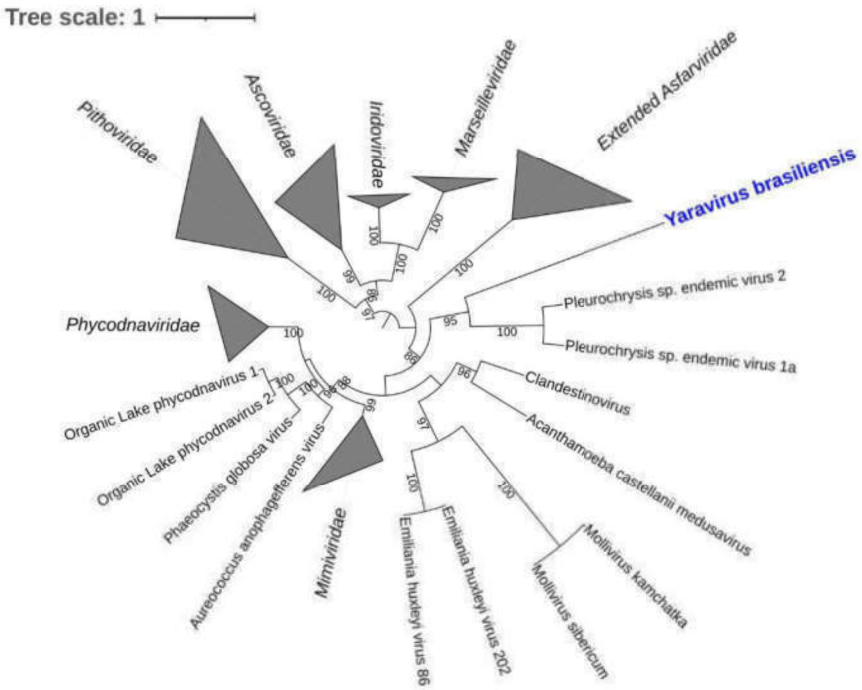


Figure 4. Maximum likelihood phylogenetic tree based on amino acid sequences of major capsid protein of Nucleocytoviricota. Yaravirus brasiliensis is bold and highlighted in blue. Sequences were aligned using Muscle [48] and low conserved regions were removed using trimAl [49]. The tree was built using IQ-TREE [50] with 1000 ultrafast bootstrap replicates and the VT+R3 model chosen by ModelTest according to Bayesian Information Criterion. The tree was visualized in iToL [51]. The tree scale indicates the substitution rate.

The genome of yaraviruses is composed mainly of ORFans, with an astonishing percentage of ~90% of their genes with functions never described before [27]. The search for yaravirus sequences in a huge dataset consisting of more than 8500 publicly available metagenomes from the most diverse habitats around the globe has also shown hits with distant homologs for the ATPase gene (NCVOG0249), with amino acid similarities that represented a number lower than 33% [27]. The discovery of yaravirus demonstrates

how important are studies focusing on isolating new viruses from the environment [27]. Although metagenomics analyses have an important role in describing new viral species by using their standard methods [67–69], microorganisms like yaraviruses would be very difficult to discover, as these protocols mostly involve recognition of genes already described [27]. This finding could also be seen as a marking point to revamp expositions in the virology field, from intriguing or persuading scientists to stimulating novel research and future researchers.

3. A Fight for Supremacy: Peculiar Features of GVs and Their Interaction with Amoeba Hosts

Much of the biology and particularities of mimivirus interactions with their hosts were discovered during early investigations of GVs. In this regard, both imaging and genomic techniques (e.g., electron microscopy, atomic force microscopy, sequencing, etc.) were of pivotal importance in uncovering many peculiar features of mimiviruses. APMV, for example, was observed to attach to the host cell through glycoside interactions between the long fibrils and surface glycans, with such adhesion also occurring with other unrelated organisms (e.g., arthropods and fungi), potentially facilitating the dispersion of these viruses in the environment [70].

Once they reach their hosts, differently from most non-giant viruses, mimiviruses (and most of the described GVs) enter the host cell through phagocytosis [71]. This was initially observed by transmission electron microscopy (TEM) analysis and further corroborated by biological assays, especially in cells treated with phagocytosis and endocytosis inhibitors [72]. At the apex of the mimivirus capsid there is a starfish-shaped protein complex that acts as a seal for the stargate, until the phagosome's internal environment promotes a new protein arrangement, unleashing the opening of the stargate and the release of the genome [73,74]. The acidification of the phagosome is suggested to be a factor that leads to capsid disassembly and membrane fusion [72,74,75].

In this regard, in 2011, the isolation of SMBV paved the way for further studies performed by other Brazilian research groups, enriching knowledge of mimiviruses' structure and biology. These studies included analysis of different Gs particles using distinct imaging techniques, such as cryo-electron microscopy (cryo-EM) and tomography, as well as fluorescence microscopy [28,76]. In a structural study developed by Schrad et al. (2020), for example, the viral particles of SMBV, tupanvirus, antarctica virus and mimivirus M4 were used to investigate the process of genome release in mimivirus-like particles [77]. Taking this work as an example, the authors have corroborated the importance of conditions such as temperature and pH for the opening of the vertex in these GVs. Here, new additional information on the viral uncoating was settled, as liberation of the viral seed (extra membrane sac) and the complete release of the viral genome were both manifested by experiments using specific conditions of these GVs' replication cycle (e.g., pH = 2 and/or 100 °C) [73]. Even though these conditions are non-biological, the authors suggest that they mimic GVs' replication cycle steps. It may also suggest that during the replication cycle other factors may play a role in capsid opening. In the same study, during the steps of viral genomic release and by adopting different imaging techniques such as cryo-EM, cryo-electron tomography and scanning electron microscopy (SEM), the authors have observed the formation of pockets devoid of DNA within the nucleocapsids of these GVs. Likewise, the analyses led to the identification of a set of proteins released from capsids during the early stages of infection within this whole complex [73]. Looking into another study, with analyses involving SEM and TEM, the authors have observed that for different mimiviruses the density of the fibrils on the surface of their capsid was variable and that this could be acquired simultaneously to genome acquisition throughout the process of morphogenesis in the large viral factories [72]. These techniques were also important in revealing key aspects of the replication cycle of different giant viruses. As already mentioned above, an antiviral strategy was beautifully described in a TEM study showing faustovirus mariensis particles trapped inside the cysts of the *Vermamoeba vermiformis* host

(Figure 5A) [26]. An in-depth description of the replication cycle of orpheovirus and cedratvirus was also established, mostly by imaging methods. For orpheovirus, we started to understand that viral exocytosis was as important as the cell lysis to the final step of this giant virus cycle (Figure 5B) [64]. For cedratvirus getuliensis, the contribution of these techniques has helped to describe a unique and complex sequential organization of the viral particle morphogenesis, including different steps of the formation of horseshoe and rectangular compartments, the incorporation of the second cork and thickening of the capsid well, and finally the formation of the ovoid-shaped virion (Figure 5C) [24]. Considering some intriguing observed features after the mimiviruses' release, another interesting study was proposed by Oliveira et al. (2019) and observed an aggregation of released tupanvirus particles with uninfected amoeba, promoting viral dissemination by the formation of host cell bunches (Figure 5D) [78]. This study revealed that this amoebal-bunch formation is correlated with the mannose-binding protein (MBP) gene expression, either induced by tupanviruses or between amoebas, through interactions among their receptor, both factors that may be important for the optimization of this process [78]. However, when we talk about the genome of tupanviruses, we observe that a great number of their genes are not present in many other mimiviruses' genomes [10], which may still hide some important information about these GVs' cycle. In view of how life has evolved on Earth, the complexity of this genome has also recently been used as an argument to suggest that viruses come from an ancestral strategy of life. According to the authors, in the period comprising the First to the Last Universal Common Ancestor (FUCA to LUCA), an intermediate ancestral (Transitional-LUCA) may have been arisen as an undifferentiated subsystem resembling a virus-like structure, from which most of the currently known viruses came [79]. Besides, aside from the already mentioned unique structural tail, and the formation of bunches [78], tupanviruses exhibited for the first time a cytotoxic phenotype to non-host cells [10]. These intriguing aspects metaphorically resemble a constant fight for supremacy [80] and help unravel the evolutionary history of GVs.

In addition to these distinct characteristics, it is worth mentioning that, as expected, the host cell does not remain indifferent to mimivirus infection. The encystment process is understood as a mechanism used by *Acanthamoeba* populations to become protected against several kinds of stressful conditions, such as dehydration, lack of nutrients, UV light, and viral infections, including against mimiviruses [26,81,82]. As observed in a study developed by Boratto et al. (2015), mimivirus infection is hampered even if those amoebas are not yet morphologically encysted but had already received the stimulus to turn into their resistant form (Figure 5E) [81]. Nonetheless, if the stimulus to become a cyst is triggered before the infection, mimiviruses as APMV are able to evade this protective status of the *Acanthamoeba* cyst, by preventing the expression of an encystment-mediating subtilisin-like serine protease and thus proceeding with the infection (Figure 5E) [81]. These studies demonstrate how complex are processes involving GVs' replication cycle and what an intricate interaction these viruses have with their amoebal hosts.

In addition to isolation studies already mentioned above, other works developed in Brazil have contributed to add knowledge of genomics and important relationships between marseilleviruses and their host. One of these studies has helped to bring light to pivotal processes in the replication cycle of marseilleviruses, specifically related to the viral entry and release. As extensively described in the literature, phagocytosis is a general route used by most GVs to enter amoebal cells [1,6,53,80]. This process is triggered only after recognition of particles larger than 500 nm [83]. However, viruses with particle sizes between 200 to 250 nm, as is the case for marseilleviruses, do not have the minimum size required to trigger this process [84].

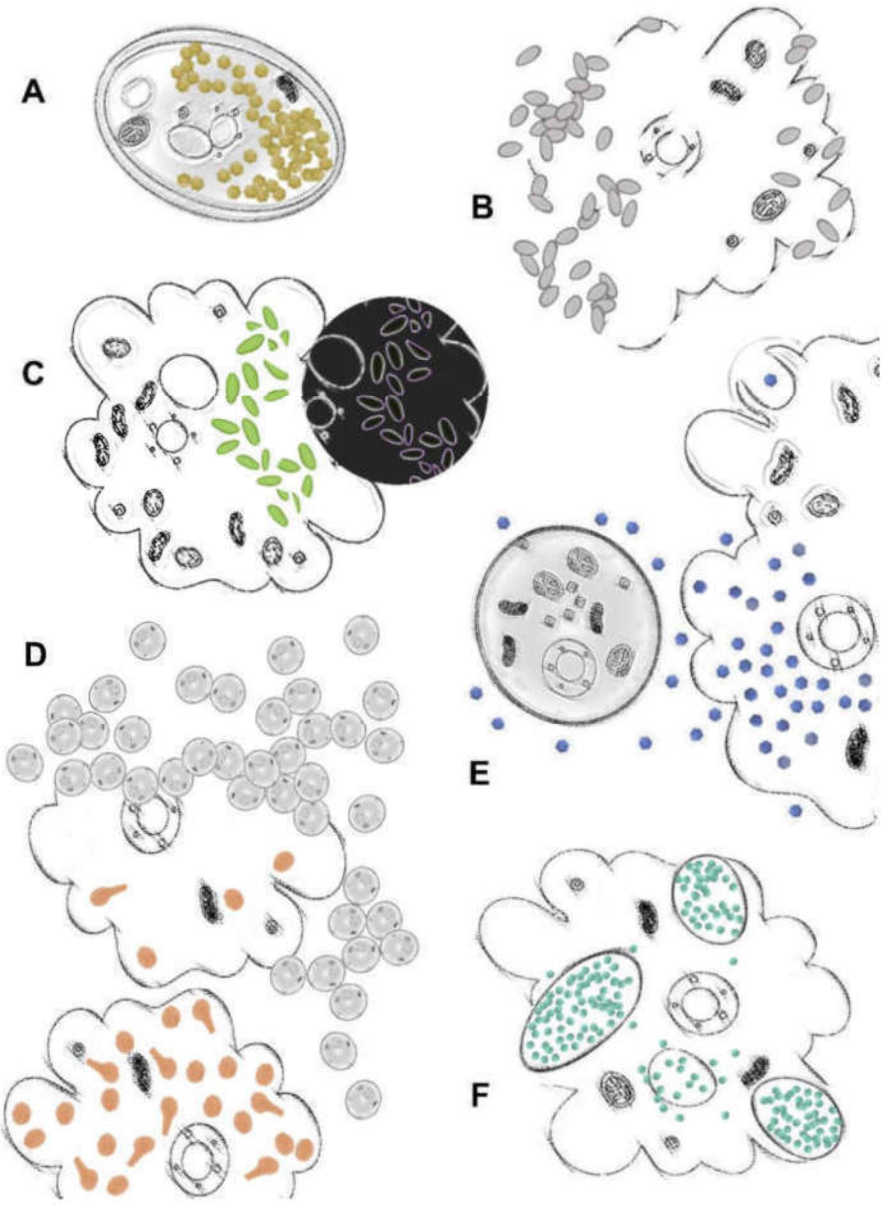


Figure 5. Unique features of giant viruses' (GVs') replication cycles unraveled in Brazil. (**A**) faustovirus dissemination is circumvented by amoebas with the enclosing of viral progeny inside the host's cysts [26]; (**B**) orpheovirus particles are released from the host by exocytosis or cell lysis [64]; (**C**) cedratvirus particles' morphogenesis follows a unique and complex sequential organization, including horseshoe and rectangular compartments, the incorporation of the second cork and thickening of the capsid well, and finally the formation of the ovoid-shaped virion [24]; (**D**) amoebas infected with tupanvirus are induced to aggregate with uninfected cells, forming giant host cell bunches [78]; (**E**) mimiviruses are able to infect amoebal trophozoites and prevent encystment, while cysts are resistant to infection [81]; (**F**) MsV are able to form giant vesicles with numerous viral particles derived from amoebal endoplasmic reticulum [84]. Amoeba images were generated from free vectors available online at Vecteezy: https://www.vecteezy.com (accessed on 22 September 2021).

By performing an in-depth investigation of the marseillevirus replication cycle and using a different set of virological assays (e.g., TEM, SEM, immunofluorescence, immunoblot-

ting), Arantes et al. have shown that during marseillevirus assembly the viral particles are organized inside large vesicles (some reaching about 3 µm in size) which are originated from the endoplasmic reticulum of the infected cells (Figure 5F) [84]. After viral release, those particles are then ready to infect another cell by exploring the phagocytosis of these vesicles that contain dozens to thousands of viral particles in their interior. In addition, viral release also seems to occur by individual virions. In this case, marseilleviruses exploit the endocytosis route to enter the cell by a mechanism which is dependent on acidification [84].

The *Marseilleviridae* family is also well known for its genomic mosaicism, which consists of the ability to incorporate foreign genes from other organisms that have *Acanthamoeba* as a common host [6]. Genomic studies of several strains of marseillevirus showed the presence of an A-T-rich promoter motif (AAATATTT) that is associated with 55% of the viral genes and that is conserved among all lineages. In addition, biological assays showed that the alteration of the promoter sequence negatively impacts the genes' transcription, showing a possible link of these sequences to the increased expression of some genes [85]. The presence of multiple copies of these motifs in the intergenic regions suggests that they may favor the fixation of newly acquired genes [85].

More recently, in 2020, analysis of the marseillevirus transcriptome revealed a temporal gene expression profile, indicating the existence of three categories: early, intermediate and late [86]. Genes belonging to different functional groups exhibited distinct expression levels throughout the infection cycle and marseillevirus infection causes significant changes in the host's transcription machinery, downregulating many genes [86].

Finally, it is worth mentioning that much of the features above described for GVs and their hosts have an influence directly affected by the intracellular environment of the amoebas. This environment has already been seen as an ecological site that comprehends a number of different and phylogenetic distant microorganisms, which not only inhabit the same location but are also observed to be in a strong process of coevolution. Even if not genetically related, an important portion of the genomic signatures (described as "the total net response to selective pressures") of the coevolving microorganisms are found to be incredibly conserved. This makes the intracellular environment of the amoebal host a sanctuary for interactions among several species of ecological and biomedical relevance [87].

4. Giant Viruses As a Tool to Update and Inspire: From the Research Fields to the Classroom

Since the known virosphere is notably anthropocentric, virology classes usually present viruses as pathogenic organisms, strongly associated with human diseases [3,88]. Instead of presenting these organisms as important tools of natural selection, ecological balance and the Earth's biogeochemical cycles, the commonly used material for teaching virology leads to a biased misconception of viruses as strictly bad, generating a certain fear in the students [89]. Besides this, other problems are the high cost of ensuring biosafety for practical virology classes, and motivation and mastery of the subject by the teachers (especially at the elementary school level) [90]. A further point is that viruses are typically very abstract for students, mainly due to their size, which limits their visualization to schematic figures, illustrations, and electron microscopy images [88].

The expansion of the perception of the virosphere by the giant viruses has unleashed a new way of understanding and teaching virology. Due to their colossal particles, the size limitation has been considered obsolete, turning these viruses into excellent learning tools [88,91]. Therefore, GVs can be visualized by optical microscopy, like bacteria and fungi, which are traditionally presented to students through common microscopes. Moreover, since they infect free-living amoebae, they represent a safe and low-cost instrument for practical virology classes [88,91].

In 2020, our group developed an educational kit to update the content typically taught in virology classes and align it to recent breakthroughs in virus research. Using slides, staining materials, viruses from the laboratory stock, and cell lineages, a microscope

slides kit called "Virus Goes Viral" was created (Figure 6A) [91]. This allows students to observe giant viruses' particles (Figure 6B–D), viral factories, and different lysis plaques in important viruses that infect animals [91]. As basic research regarding GVs in the Brazilian territory may enrich the knowledge of these microorganisms in the virology field, this kit may also aim to foster an inspiring learning environment, as well as ignite more interest in these fascinating organisms in the future.

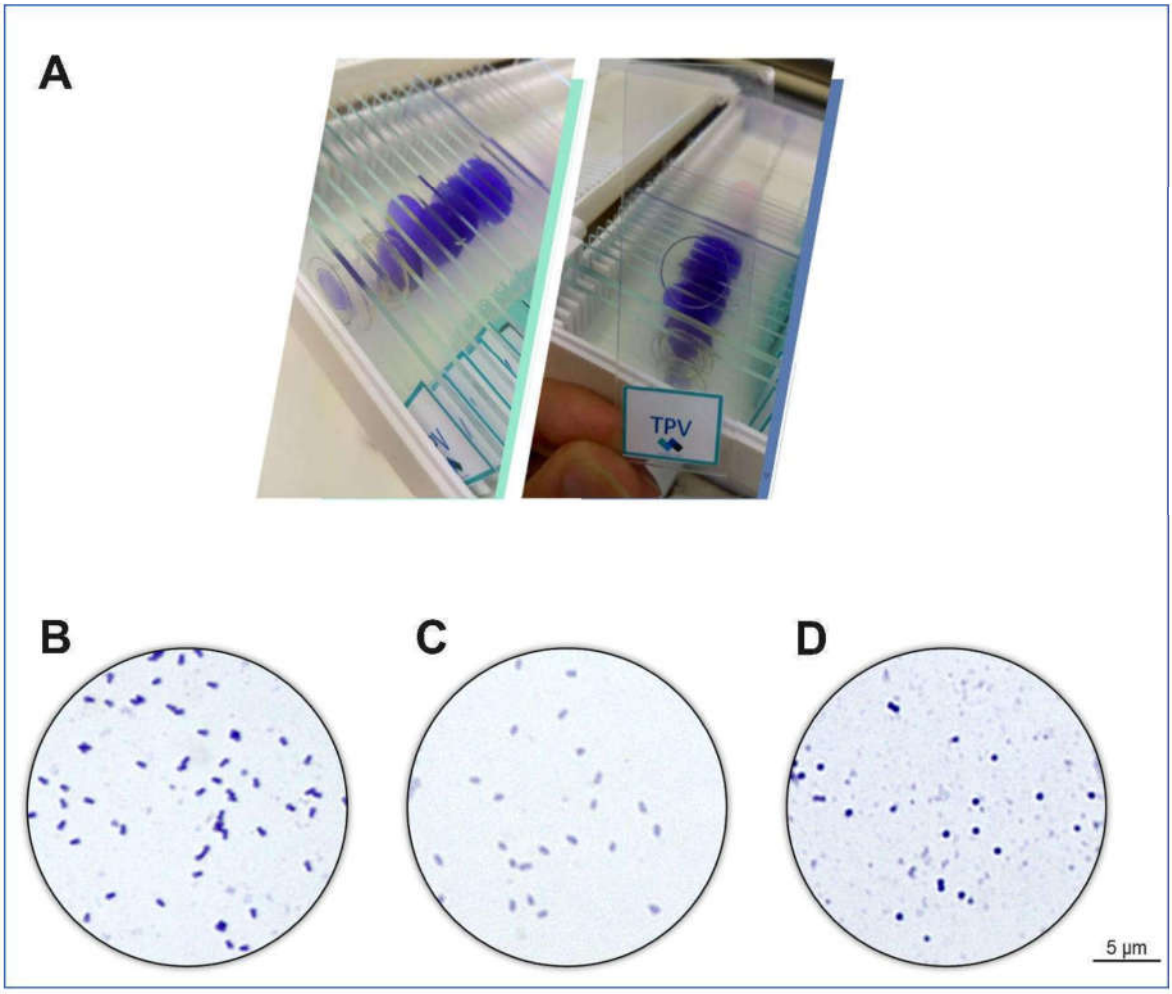


Figure 6. Optical microscopy images. (**A**) Source: Reference [91]. Visualization under 1000 times magnification of the stained purified particles of (**B**) tupanvirus, (**C**) cedratvirus and (**D**) Niemeyer virus, respectively.

5. Conclusions

The serendipitous discovery of APMV in 2003 changed the concept of viruses and expanded the limits of the virosphere [1]. Over the last two decades, different groups of large and giant viruses of amoebae have been described throughout the world, revealing many unusual particles' shapes and genome length and content. Culture-independent studies have proved the ubiquity and the astonishing diversity of giant viruses on Earth [18,68]. Now, we must go deeper into the characterization of these viruses, by using different isolates as models. Several new viruses isolated from distinct viral families (e.g., *Mimiviridae* and *Marseilleviridae*) are now under in-depth investigation to better understand their biology and evolution, and these outstanding discoveries may even change our perception of

life itself (e.g., *Mimiviridae* as a new branch derived from a population that gave origin to the modern Eukarya [20]).

Brazil is the fifth largest country in terms of territory and harbors different biomes, making the country an important global hub of tropical biodiversity. Over the past 10 years, the diversity of amoebal viruses in Brazilian environments has been uncovered, with hundreds of isolates belonging to distinct groups, including members of *Mimiviridae*, *Marseilleviridae*, Pandoraviridae, Pithoviridae, Faustoviridae, *Lavidaviridae*, among others, as discussed in previous sections. Interestingly, the most complex giant viruses described to date were isolated from two distinct areas in Brazil [10]. Furthermore, the recent discovery of the small mysterious yaravirus in the country highlighted the importance of continuing the search for new isolates, which could reveal completely new entities on Earth [27]. Brazilian groups, working alongside other experts in the field, have contributed to uncovering this unusual and exciting side of the virosphere. The study of amoebal viruses has already changed our perception of basic virology. Furthermore, giant viruses have recently been proposed as tools to improve virology learning at different educational levels [91]. Surely, other potential applications for these viruses are waiting to be revealed as new data emerges. This is an open field for remarkable discoveries, and we can expect great innovations as new amoebal viruses are isolated and characterized. In this context, the preservation of Brazilian biomes is a sine qua non condition, not only for the discovery of novel biological entities (including giant viruses), but also because of climatic, philosophical, political and economic reasons. Finally, we would like to reinforce that, although this review is a celebration of the 10th anniversary of giant virus studies in Brazilian biomes, we do not support any kind of excessive scientific nationalism. We are aware that studies of giant viruses in Brazil represent a modest contribution to the giant viruses' universe. This field has been constructed by remarkable worldwide research groups, and we are very grateful for their efforts and inspiring work.

Author Contributions: Conceptualization, P.V.M.B., M.S.M.S., A.S.A.W., A.P.C.C., B.L.d.A., G.A.P.d.S., I.L.M.d.A., T.B.M., V.F.Q., R.A.L.R. Writing—original draft preparation, P.V.M.B., M.S.M.S., A.S.A.W., A.P.C.C., B.L.d.A., G.A.P.d.S., I.L.M.d.A., T.B.M., V.F.Q., R.A.L.R. Writing—review and editing, P.V.M.B., M.S.M.S., I.B., J.R.C., S.T.d.F., R.N.d.S., F.S.C., A.C.F., J.S.A. Supervision and funding acquisition, J.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: JSA is a CNPq researcher (302081/2018-6).

Institutional Review Board Statement: The data here presented was partially or fully registered at SISGEN—numbers AA3B21E, A702EB8, A25764F, AC31840, A473BD3, A3DAB3F, AC3045D, A96431C, ABF23CC, A2F8816, A580BBD, AEC3EAA. Collection authorization -SISBIO numbers 33326, 34293 and 80252.

Informed Consent Statement: Not applicable.

Data Availability Statement: Genomic data can be found at genbank (https://www.ncbi.nlm.nih.gov/genbank/, accessed on 11 January 2022).

Acknowledgments: We would like to thank all colleagues that contributed to studies on giant viruses in Brazil and worldwide. It has been a long, hard, challenging but pleasant journey. We are grateful to Bernard La Scola (Aix Marseille Université, France) for those years of fruitful collaboration. We also thank the institutes that provided support in the last 10 years, including the CNPq, CAPES, FAPEMIG, Ministério da Saúde, Ministério de Meio Ambiente, Ministério da Educação and Pró-Reitoria de Pesquisa e de Pós-Graduacão da UFMG and Centro de Microscopia da UFMG.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Scola, B.L.; 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. [CrossRef]
2. Iyer, L.M.; Aravind, L.; Koonin, E.V. Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. *J. Virol.* **2001**, *75*, 11720–11734. [CrossRef] [PubMed]

3. Rodrigues, R.A.; Andrade, A.C.; Boratto, P.V.d.M.; Trindade, G.d.S.; Kroon, E.G.; Abrahão, J.S. An Anthropocentric View of the Virosphere-Host Relationship. *Front. Microbiol.* **2017**, *8*, 1673. [CrossRef] [PubMed]
4. La Scola, B.; Desnues, C.; Pagnier, I.; Robert, C.; Barrassi, L.; Fournous, G.; Merchat, M.; Suzan-Monti, M.; Forterre, P.; Koonin, E.; et al. The Virophage as a Unique Parasite of the Giant Mimivirus. *Nature* **2008**, *455*, 100–104. [CrossRef] [PubMed]
5. Koonin, E.V.; Yutin, N. Origin and Evolution of Eukaryotic Large Nucleo-Cytoplasmic DNA Viruses. *Intervirology* **2010**, *53*, 284–292. [CrossRef]
6. Boyer, M.; Yutin, N.; Pagnier, I.; Barrassi, L.; Fournous, G.; Espinosa, L.; Robert, C.; Azza, S.; Sun, S.; Rossmann, M.G.; et al. Giant Marseillevirus Highlights the Role of Amoebae as a Melting Pot in Emergence of Chimeric Microorganisms. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21848–21853. [CrossRef] [PubMed]
7. Andrade, A.C.D.S.P.; Arantes, T.S.; Rodrigues, R.A.L.; Machado, T.B.; Dornas, F.P.; Landell, M.F.; Furst, C.; Borges, L.G.A.; Dutra, L.A.L.; Almeida, G.; et al. Ubiquitous Giants: A Plethora of Giant Viruses Found in Brazil and Antarctica. *Virol. J.* **2018**, *15*, 22. [CrossRef]
8. Aherfi, S.; Colson, P.; La Scola, B.; Raoult, D. Giant Viruses of Amoebas: An Update. *Front. Microbiol.* **2016**, *7*, 349. [CrossRef]
9. Schulz, F.; Yutin, N.; Ivanova, N.N.; Ortega, D.R.; Lee, T.K.; Vierheilig, J.; Daims, H.; Horn, M.; Wagner, M.; Jensen, G.J.; et al. Giant Viruses with an Expanded Complement of Translation System Components. *Science* **2017**, *356*, 82–85. [CrossRef]
10. Abrahão, J.; Silva, L.; Silva, L.S.; Khalil, J.Y.B.; Rodrigues, R.; Arantes, T.; Assis, F.; Boratto, P.; Andrade, M.; Kroon, E.G.; et al. Tailed Giant Tupanvirus Possesses the Most Complete Translational Apparatus of the Known Virosphere. *Nat. Commun.* **2018**, *9*, 749. [CrossRef] [PubMed]
11. Yoshikawa, G.; Blanc-Mathieu, R.; Song, C.; Kayama, Y.; Mochizuki, T.; Murata, K.; Ogata, H.; Takemura, M. Medusavirus, a Novel Large DNA Virus Discovered from Hot Spring Water. *J. Virol.* **2019**, *93*, e02130-18. [CrossRef] [PubMed]
12. 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. USA* **2015**, *112*, E5327–E5335. [CrossRef] [PubMed]
13. Abrahão, J.S.; Dornas, F.P.; Silva, L.C.; Almeida, G.M.; Boratto, P.V.; Colson, P.; La Scola, B.; Kroon, E.G. Acanthamoeba Polyphaga Mimivirus and Other Giant Viruses: An Open Field to Outstanding Discoveries. *Virol. J.* **2014**, *11*, 120. [CrossRef] [PubMed]
14. Marciano-Cabral, F.; Cabral, G. *Acanthamoeba* spp. as Agents of Disease in Humans. *Clin. Microbiol. Rev.* **2003**, *16*, 273–307. [CrossRef] [PubMed]
15. Colson, P.; La Scola, B.; Raoult, D. Giant Viruses of Amoebae: A Journey through Innovative Research and Paradigm Changes. *Annu. Rev. Virol.* **2017**, *4*, 61–85. [CrossRef]
16. Andrade, K.R.; Boratto, P.P.V.M.; Rodrigues, F.P.; Silva, L.C.F.; Dornas, F.P.; Pilotto, M.R.; La Scola, B.; Almeida, G.M.F.; Kroon, E.G.; Abrahão, J.S. Oysters as Hot Spots for Mimivirus Isolation. *Arch. Virol.* **2015**, *160*, 477–482. [CrossRef]
17. Mihara, T.; Koyano, H.; Hingamp, P.; Grimsley, N.; Goto, S.; Ogata, H. Taxon Richness of "Megaviridae" Exceeds Those of Bacteria and Archaea in the Ocean. *Microbes Environ.* **2018**, *33*, 162–171. [CrossRef] [PubMed]
18. 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.* **2020**, *11*, 1710. [CrossRef]
19. Moniruzzaman, M.; Weinheimer, A.R.; Martinez-Gutierrez, C.A.; Aylward, F.O. Widespread Endogenization of Giant Viruses Shapes Genomes of Green Algae. *Nature* **2020**, *588*, 141–145. [CrossRef]
20. Marcelino, V.M.; Espinola, M.V.P.C.; Serrano-Solis, V.; Farias, S.T. Evolution of the Genus Mimivirus Based on Translation Protein Homology and Its Implication in the Tree of Life. *Genet. Mol. Res.* **2017**, *16*, 1–7. [CrossRef]
21. Dornas, F.P.; Khalil, J.Y.B.; Pagnier, I.; Raoult, D.; Abrahão, J.; La Scola, B. Isolation of New Brazilian Giant Viruses from Environmental Samples Using a Panel of Protozoa. *Front. Microbiol.* **2015**, *6*, 1086. [CrossRef] [PubMed]
22. Campos, R.K.; Boratto, P.V.; Assis, F.L.; Aguiar, E.R.; Silva, L.C.; Albarnaz, J.D.; Dornas, F.P.; Trindade, G.S.; Ferreira, P.P.; Marques, J.T.; et al. Samba Virus: A Novel Mimivirus from a Giant Rain Forest, the Brazilian Amazon. *Virol. J.* **2014**, *11*, 95. [CrossRef]
23. Guerreiro, R.L.; Bergier, I.; McGlue, M.M.; Warren, L.V.; de Abreu, U.G.P.; Abrahão, J.; Assine, M.L. The Soda Lakes of Nhecolândia: A Conservation Opportunity for the Pantanal Wetlands. *Perspect. Ecol. Conserv.* **2019**, *17*, 9–18. [CrossRef]
24. Dos Santos Silva, L.K.; Andrade, A.C.S.P.; Dornas, F.P.; Rodrigues, R.A.L.; Arantes, T.; Kroon, E.G.; Bonjardim, C.A.; Abrahão, J.S. Cedratvirus Getuliensis Replication Cycle: An in-Depth Morphological Analysis. *Sci. Rep.* **2018**, *8*, 4000. [CrossRef]
25. Boratto, P.V.M.; Arantes, T.S.; Silva, L.C.F.; Assis, F.L.; Kroon, E.G.; La Scola, B.; Abrahão, J.S. Niemeyer Virus: A New Mimivirus Group A Isolate Harboring a Set of Duplicated Aminoacyl-TRNA Synthetase Genes. *Front. Microbiol.* **2015**, *6*, 1256. [CrossRef]
26. Borges, I.; Rodrigues, R.A.L.; Dornas, F.P.; Almeida, G.; Aquino, I.; Bonjardim, C.A.; Kroon, E.G.; La Scola, B.; Abrahão, J.S. Trapping the Enemy: *Vermamoeba vermiformis* Circumvents Faustovirus Mariensis Dissemination by Enclosing Viral Progeny inside Cysts. *J. Virol.* **2019**, *93*, e00312-19. [CrossRef]
27. Boratto, P.V.M.; Oliveira, G.P.; Machado, T.B.; Andrade, A.C.S.P.; Baudoin, J.-P.; Klose, T.; Schulz, F.; Azza, S.; Decloquement, P.; Chabrière, E.; et al. Yaravirus: A Novel 80-Nm Virus Infecting Acanthamoeba Castellanii. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 16579–16586. [CrossRef]
28. Schrad, J.R.; Young, E.J.; Abrahão, J.S.; Cortines, J.R.; Parent, K.N. Microscopic Characterization of the Brazilian Giant Samba Virus. *Viruses* **2017**, *9*, 30. [CrossRef] [PubMed]

29. Borges, I.A.; de Assis, F.L.; Silva, L.K.D.S.; Abrahão, J. Rio Negro Virophage: Sequencing of the near Complete Genome and Transmission Electron Microscopy of Viral Factories and Particles. *Braz. J. Microbiol.* **2018**, *49* (Suppl. S1), 260–261. [CrossRef] [PubMed]
30. Mougari, S.; Bekliz, M.; Abrahao, J.; Di Pinto, F.; Levasseur, A.; La Scola, B. Guarani Virophage, a New Sputnik-Like Isolate From a Brazilian Lake. *Front. Microbiol.* **2019**, *10*, 1003. [CrossRef]
31. Assis, F.L.; Franco-Luiz, A.P.M.; Dos Santos, R.N.; Campos, F.S.; Dornas, F.P.; Boratto, P.V.M.; Franco, A.C.; Abrahao, J.S.; Colson, P.; Scola, B.L. Genome Characterization of the First Mimiviruses of Lineage C Isolated in Brazil. *Front. Microbiol.* **2017**, *8*, 2562. [CrossRef] [PubMed]
32. Boratto, P.V.M.; Dornas, F.P.; da Silva, L.C.F.; Rodrigues, R.A.L.; Oliveira, G.P.; Cortines, J.R.; Drumond, B.P.; Abrahão, J.S. Analyses of the Kroon Virus Major Capsid Gene and Its Transcript Highlight a Distinct Pattern of Gene Evolution and Splicing among Mimiviruses. *J. Virol.* **2018**, *92*, e01782-17. [CrossRef]
33. Assis, F.L.; Bajrai, L.; Abrahao, J.S.; Kroon, E.G.; Dornas, F.P.; Andrade, K.R.; Boratto, P.V.M.; Pilotto, M.R.; Robert, C.; Benamar, S.; et al. Pan-Genome Analysis of Brazilian Lineage A Amoebal Mimiviruses. *Viruses* **2015**, *7*, 3483–3499. [CrossRef] [PubMed]
34. Rodrigues, R.A.L.; Mougari, S.; Colson, P.; La Scola, B.; Abrahão, J.S. "Tupanvirus", a New Genus in the Family Mimiviridae. *Arch. Virol.* **2019**, *164*, 325–331. [CrossRef]
35. De Miranda Boratto, P.V.; Dos Santos Pereira Andrade, A.C.; Araújo Lima Rodrigues, R.; La Scola, B.; Santos Abrahão, J. The Multiple Origins of Proteins Present in Tupanvirus Particles. *Curr. Opin. Virol.* **2019**, *36*, 25–31. [CrossRef] [PubMed]
36. Raoult, D. The Post-Darwinist Rhizome of Life. *Lancet* **2010**, *375*, 104–105. [CrossRef]
37. Abrahão, J.S.; Araújo, R.; Colson, P.; Scola, B.L. The Analysis of Translation-Related Gene Set Boosts Debates around Origin and Evolution of Mimiviruses. *PLoS Genet.* **2017**, *13*, e1006532. [CrossRef]
38. Aherfi, S.; La Scola, B.; Pagnier, I.; Raoult, D.; Colson, P. The Expanding Family Marseilleviridae. *Virology* **2014**, *466–467*, 27–37. [CrossRef]
39. Thomas, V.; Bertelli, C.; Collyn, F.; Casson, N.; Telenti, A.; Goesmann, A.; Croxatto, A.; Greub, G. Lausannevirus, a Giant Amoebal Virus Encoding Histone Doublets. *Environ. Microbiol.* **2011**, *13*, 1454–1466. [CrossRef]
40. Aherfi, S.; Pagnier, I.; Fournous, G.; Raoult, D.; La Scola, B.; Colson, P. Complete Genome Sequence of Cannes 8 Virus, a New Member of the Proposed Family "Marseilleviridae". *Virus Genes* **2013**, *47*, 550–555. [CrossRef]
41. Pagnier, I.; Reteno, D.-G.I.; Saadi, H.; Boughalmi, M.; Gaia, M.; Slimani, M.; Ngounga, T.; Bekliz, M.; Colson, P.; Raoult, D.; et al. A Decade of Improvements in Mimiviridae and Marseilleviridae Isolation from Amoeba. *Intervirology* **2013**, *56*, 354–363. [CrossRef]
42. Aherfi, S.; Boughalmi, M.; Pagnier, I.; Fournous, G.; La Scola, B.; Raoult, D.; Colson, P. Complete Genome Sequence of Tunisvirus, a New Member of the Proposed Family Marseilleviridae. *Arch. Virol.* **2014**, *159*, 2349–2358. [CrossRef] [PubMed]
43. Boughalmi, M.; Pagnier, I.; Aherfi, S.; Colson, P.; Raoult, D.; La Scola, B. First Isolation of a Marseillevirus in the Diptera Syrphidae Eristalis Tenax. *Intervirology* **2013**, *56*, 386–394. [CrossRef]
44. Lagier, J.-C.; Armougom, F.; Million, M.; Hugon, P.; Pagnier, I.; Robert, C.; Bittar, F.; Fournous, G.; Gimenez, G.; Maraninchi, M.; et al. Microbial Culturomics: Paradigm Shift in the Human Gut Microbiome Study. *Clin. Microbiol. Infect.* **2012**, *18*, 1185–1193. [CrossRef]
45. Colson, P.; Fancello, L.; Gimenez, G.; Armougom, F.; Desnues, C.; Fournous, G.; Yoosuf, N.; Million, M.; La Scola, B.; Raoult, D. Evidence of the Megavirome in Humans. *J. Clin. Virol.* **2013**, *57*, 191–200. [CrossRef] [PubMed]
46. Doutre, G.; Philippe, N.; Abergel, C.; Claverie, J.-M. Genome Analysis of the First Marseilleviridae Representative from Australia Indicates That Most of Its Genes Contribute to Virus Fitness. *J. Virol.* **2014**, *88*, 14340–14349. [CrossRef]
47. Dornas, F.P.; Assis, F.L.; Aherfi, S.; Arantes, T.; Abrahão, J.S.; Colson, P.; La Scola, B. A Brazilian Marseillevirus Is the Founding Member of a Lineage in Family Marseilleviridae. *Viruses* **2016**, *8*, 76. [CrossRef] [PubMed]
48. Edgar, R.C. MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity. *BMC Bioinform.* **2004**, *5*, 113. [CrossRef] [PubMed]
49. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]
50. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]
51. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]
52. Dos Santos, R.N.; Campos, F.S.; Medeiros de Albuquerque, N.R.; Finoketti, F.; Côrrea, R.A.; Cano-Ortiz, L.; Assis, F.L.; Arantes, T.S.; Roehe, P.M.; Franco, A.C. A New Marseillevirus Isolated in Southern Brazil from Limnoperna Fortunei. *Sci. Rep.* **2016**, *6*, 35237. [CrossRef]
53. 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. [CrossRef]
54. Scheid, P.; Balczun, C.; Schaub, G.A. Some Secrets Are Revealed: Parasitic Keratitis Amoebae as Vectors of the Scarcely Described Pandoraviruses to Humans. *Parasitol. Res.* **2014**, *113*, 3759–3764. [CrossRef] [PubMed]

55. Antwerpen, M.H.; Georgi, E.; Zoeller, L.; Woelfel, R.; Stoecker, K.; Scheid, P. Whole-Genome Sequencing of a Pandoravirus Isolated from Keratitis-Inducing Acanthamoeba. *Genome Announc.* **2015**, *3*, e00136-15. [CrossRef]
56. Legendre, M.; Alempic, J.-M.; Philippe, N.; Lartigue, A.; Jeudy, S.; Poirot, O.; Ta, N.T.; Nin, S.; Couté, Y.; Abergel, C.; et al. Pandoravirus Celtis Illustrates the Microevolution Processes at Work in the Giant Pandoraviridae Genomes. *Front. Microbiol.* **2019**, *10*, 430. [CrossRef]
57. Legendre, M.; Fabre, E.; Poirot, O.; Jeudy, S.; Lartigue, A.; Alempic, J.-M.; Beucher, L.; Philippe, N.; Bertaux, L.; Christo-Foroux, E.; et al. Diversity and Evolution of the Emerging Pandoraviridae Family. *Nat. Commun.* **2018**, *9*, 2285. [CrossRef]
58. Pereira Andrade, A.C.D.S.; Victor de Miranda Boratto, P.; Rodrigues, R.A.L.; Bastos, T.M.; Azevedo, B.L.; Dornas, F.P.; Oliveira, D.B.; Drumond, B.P.; Kroon, E.G.; Abrahão, J.S. New Isolates of Pandoraviruses: Contribution to the Study of Replication Cycle Steps. *J. Virol.* **2019**, *93*, e01942-18. [CrossRef] [PubMed]
59. Andreani, J.; Aherfi, S.; Bou Khalil, J.Y.; 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*, 300. [CrossRef]
60. Bertelli, C.; Mueller, L.; Thomas, V.; Pillonel, T.; Jacquier, N.; Greub, G. Cedratvirus Lausannensis—Digging into Pithoviridae Diversity. *Environ. Microbiol.* **2017**, *19*, 4022–4034. [CrossRef]
61. Jeudy, S.; Rigou, S.; Alempic, J.-M.; Claverie, J.-M.; Abergel, C.; Legendre, M. The DNA Methylation Landscape of Giant Viruses. *Nat. Commun.* **2020**, *11*, 2657. [CrossRef]
62. Andreani, J.; Khalil, J.Y.B.; 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*, 2643. [CrossRef]
63. Rodrigues, R.A.L.; Andreani, J.; Andrade, A.C.D.S.P.; Machado, T.B.; Abdi, S.; Levasseur, A.; Abrahão, J.S.; La Scola, B. Morphologic and Genomic Analyses of New Isolates Reveal a Second Lineage of Cedratviruses. *J. Virol.* **2018**, *92*, e00372-18. [CrossRef]
64. Souza, F.; Rodrigues, R.; Reis, E.; Lima, M.; La Scola, B.; Abrahão, J. In-Depth Analysis of the Replication Cycle of Orpheovirus. *Virol. J.* **2019**, *16*, 158. [CrossRef]
65. Reteno, D.G.; Benamar, S.; Khalil, J.B.; 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. [CrossRef] [PubMed]
66. Oliveira, G.P.; de Aquino, I.L.M.; Luiz, A.P.M.F.; Abrahão, J.S. Putative Promoter Motif Analyses Reinforce the Evolutionary Relationships Among Faustoviruses, Kaumoebavirus, and Asfarvirus. *Front. Microbiol.* **2018**, *9*, 1041. [CrossRef] [PubMed]
67. Tully, B.J.; Graham, E.D.; Heidelberg, J.F. The Reconstruction of 2,631 Draft Metagenome-Assembled Genomes from the Global Oceans. *Sci. Data* **2018**, *5*, 170203. [CrossRef]
68. Schulz, F.; Roux, S.; Paez-Espino, D.; Jungbluth, S.; Walsh, D.A.; Denef, V.J.; McMahon, K.D.; Konstantinidis, K.T.; Eloe-Fadrosh, E.A.; Kyrpides, N.C.; et al. Giant Virus Diversity and Host Interactions through Global Metagenomics. *Nature* **2020**, *578*, 432–436. [CrossRef]
69. Hingamp, P.; Grimsley, N.; Acinas, S.G.; Clerissi, C.; Subirana, L.; Poulain, J.; Ferrera, I.; Sarmento, H.; Villar, E.; Lima-Mendez, G.; et al. Exploring Nucleo-Cytoplasmic Large DNA Viruses in Tara Oceans Microbial Metagenomes. *ISME J.* **2013**, *7*, 1678–1695. [CrossRef] [PubMed]
70. Rodrigues, R.A.L.; dos Santos Silva, L.K.; Dornas, F.P.; de Oliveira, D.B.; Magalhães, T.F.F.; Santos, D.A.; Costa, A.O.; de Macêdo Farias, L.; Magalhães, P.P.; Bonjardim, C.A.; et al. Mimivirus Fibrils Are Important for Viral Attachment to the Microbial World by a Diverse Glycoside Interaction Repertoire. *J. Virol.* **2015**, *89*, 11812–11819. [CrossRef]
71. Ghigo, E.; Kartenbeck, J.; Lien, P.; Pelkmans, L.; Capo, C.; Mege, J.-L.; Raoult, D. Ameobal Pathogen Mimivirus Infects Macrophages through Phagocytosis. *PLoS Pathog.* **2008**, *4*, e1000087. [CrossRef]
72. Andrade, A.C.D.S.P.; Rodrigues, R.A.L.; Oliveira, G.P.; Andrade, K.R.; Bonjardim, C.A.; La Scola, B.; Kroon, E.G.; Abrahão, J.S. Filling Knowledge Gaps for Mimivirus Entry, Uncoating, and Morphogenesis. *J. Virol.* **2017**, *91*, e01335-17. [CrossRef] [PubMed]
73. Schrad, J.R.; Abrahão, J.S.; Cortines, J.R.; Parent, K.N. Structural and Proteomic Characterization of the Initiation of Giant Virus Infection. *Cell* **2020**, *181*, 1046–1061.e6. [CrossRef] [PubMed]
74. De Souza, G.A.P.; Queiroz, V.F.; Coelho, L.F.L.; Abrahão, J.S. Alohomora! What the Entry Mechanisms Tell Us about the Evolution and Diversification of Giant Viruses and Their Hosts. *Curr. Opin. Virol.* **2021**, *47*, 79–85. [CrossRef] [PubMed]
75. Quemin, E.R.; Corroyer-Dulmont, S.; Krijnse-Locker, J. Entry and Disassembly of Large DNA Viruses: Electron Microscopy Leads the Way. *J. Mol. Biol.* **2018**, *430*, 1714–1724. [CrossRef]
76. Parent, K.N.; Schrad, J.R.; Young, E.J.; Abrahao, J.S.; Cortines, J.R. A Gateway into Understanding the Unique Vertex of Samba Virus. *Microsc. Microanal.* **2018**, *24*, 1438–1439. [CrossRef]
77. Zauberman, N.; Mutsafi, Y.; Halevy, D.B.; 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*, e114. [CrossRef] [PubMed]
78. Oliveira, G.; Silva, L.; Leão, T.; Mougari, S.; da Fonseca, F.G.; Kroon, E.G.; La Scola, B.; Abrahão, J.S. Tupanvirus-Infected Amoebas Are Induced to Aggregate with Uninfected Cells Promoting Viral Dissemination. *Sci. Rep.* **2019**, *9*, 183. [CrossRef] [PubMed]
79. De Farias, S.T.; Jheeta, S.; Prosdocimi, F. Viruses as a Survival Strategy in the Armory of Life. *Hist. Philos. Life Sci.* **2019**, *41*, 45. [CrossRef] [PubMed]
80. Oliveira, G.; La Scola, B.; Abrahão, J. Giant Virus vs Amoeba: Fight for Supremacy. *Virol. J.* **2019**, *16*, 126. [CrossRef]

81. Boratto, P.; Albarnaz, J.D.; Almeida, G.M.; Botelho, L.; Fontes, A.C.L.; Costa, A.O.; Santos, D.d.A.; Bonjardim, C.A.; La Scola, B.; Kroon, E.G.; et al. Acanthamoeba Polyphaga Mimivirus Prevents Amoebal Encystment-Mediating Serine Proteinase Expression and Circumvents Cell Encystment. *J. Virol.* **2015**, *89*, 2962–2965. [CrossRef]
82. Silva, L.K.D.S.; Boratto, P.V.M.; La Scola, B.; Bonjardim, C.A.; Abrahão, J.S. Acanthamoeba and Mimivirus Interactions: The Role of Amoebal Encystment and the Expansion of the "Cheshire Cat" Theory. *Curr. Opin. Microbiol.* **2016**, *31*, 9–15. [CrossRef] [PubMed]
83. Korn, E.D.; Weisman, R.A. Phagocytosis of Latex Beads by Acanthamoeba. II. Electron Microscopic Study of the Initial Events. *J. Cell Biol.* **1967**, *34*, 219–227. [CrossRef] [PubMed]
84. Arantes, T.S.; Rodrigues, R.A.L.; Dos Santos Silva, L.K.; Oliveira, G.P.; de Souza, H.L.; Khalil, J.Y.B.; de Oliveira, D.B.; Torres, A.A.; da Silva, L.L.; Colson, P.; et al. The Large Marseillevirus Explores Different Entry Pathways by Forming Giant Infectious Vesicles. *J. Virol.* **2016**, *90*, 5246–5255. [CrossRef]
85. Oliveira, G.P.; Lima, M.T.; Arantes, T.S.; Assis, F.L.; Rodrigues, R.A.L.; da Fonseca, F.G.; Bonjardim, C.A.; Kroon, E.G.; Colson, P.; La Scola, B.; et al. The Investigation of Promoter Sequences of Marseilleviruses Highlights a Remarkable Abundance of the AAATATTT Motif in Intergenic Regions. *J. Virol.* **2017**, *91*, e01088-17. [CrossRef] [PubMed]
86. Rodrigues, R.A.L.; Louazani, A.C.; Picorelli, A.; Oliveira, G.P.; Lobo, F.P.; Colson, P.; La Scola, B.; Abrahão, J.S. Analysis of a Marseillevirus Transcriptome Reveals Temporal Gene Expression Profile and Host Transcriptional Shift. *Front. Microbiol.* **2020**, *11*, 651. [CrossRef]
87. Serrano-Solís, V.; Toscano Soares, P.E.; de Farías, S.T. Genomic Signatures Among Acanthamoeba Polyphaga Entoorganisms Unveil Evidence of Coevolution. *J. Mol. Evol.* **2019**, *87*, 7–15. [CrossRef]
88. Akashi, M.; Fukaya, S.; Uchiyama, C.; Aoki, K.; Takemura, M. Visualization of Giant Virus Particles and Development of "VIRAMOS" for High School and University Biology Course. *Biochem. Mol. Biol. Educ.* **2019**, *47*, 426–431. [CrossRef]
89. Paez-Espino, D.; Eloe-Fadrosh, E.A.; Pavlopoulos, G.A.; Thomas, A.D.; Huntemann, M.; Mikhailova, N.; Rubin, E.; Ivanova, N.N.; Kyrpides, N.C. Uncovering Earth's Virome. *Nature* **2016**, *536*, 425–430. [CrossRef]
90. Matza-Porges, S.; Nathan, D. A Biosafety Level 2 Virology Lab for Biotechnology Undergraduates. *Biochem. Mol. Biol. Educ.* **2017**, *45*, 537–543. [CrossRef]
91. De Souza, G.A.P.; Queiroz, V.F.; Lima, M.T.; de Sousa Reis, E.V.; Coelho, L.F.L.; Abrahão, J.S. Virus Goes Viral: An Educational Kit for Virology Classes. *Virol. J.* **2020**, *17*, 13. [CrossRef] [PubMed]

6. Outros artigos:

Nas páginas a seguir estão anexados os artigos dos quais pude participar como coautor durante esse período de doutorado. Segue a lista de artigos:

6.1 *Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere.*

6.2 *Single Cell Micro-aspiration as an Alternative Strategy to Fluorescence-activated Cell Sorting for Giant Virus Mixture Separation.*

6.3 *Detection of SARS-CoV-2 RNA on public surfaces in a densely populated urban area of Brazil*

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-03168-1 OPEN

Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere

Jônatas Abrahão[1,2], Lorena Silva[1,2], Ludmila Santos Silva[1,2], Jacques Yaacoub Bou Khalil[3], Rodrigo Rodrigues[2], Thalita Arantes[2], Felipe Assis[2], Paulo Boratto[2], Miguel Andrade[4], Erna Geessien Kroon[2], Bergmann Ribeiro[4], Ivan Bergier[5], Herve Seligmann[1], Eric Ghigo[1], Philippe Colson[1], Anthony Levasseur[1], Guido Kroemer[6,7,8,9,10,11,12], Didier Raoult[1] & Bernard La Scola[1]

Here we report the discovery of two Tupanvirus strains, the longest tailed *Mimiviridae* members isolated in amoebae. Their genomes are 1.44–1.51 Mb linear double-strand DNA coding for 1276–1425 predicted proteins. Tupanviruses share the same ancestors with mimivirus lineages and these giant viruses present the largest translational apparatus within the known virosphere, with up to 70 tRNA, 20 aaRS, 11 factors for all translation steps, and factors related to tRNA/mRNA maturation and ribosome protein modification. Moreover, two sequences with significant similarity to intronic regions of 18 S rRNA genes are encoded by the tupanviruses and highly expressed. In this translation-associated gene set, only the ribosome is lacking. At high multiplicity of infections, tupanvirus is also cytotoxic and causes a severe shutdown of ribosomal RNA and a progressive degradation of the nucleus in host and non-host cells. The analysis of tupanviruses constitutes a new step toward understanding the evolution of giant viruses.

[1] MEPHI, APHM, IRD 198, Aix Marseille Univ, IHU-Méditerranee Infection, 19-21 Bd Jean Moulin, 13005 Marseille, France. [2] Laboratório de Vírus, Instituto de Ciências Biológicas, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil. [3] CNRS, 13005 Marseille, France. [4] Laboratório de Microscopia Eletrônica e Virologia, Departamento de Biologia Celular, Instituto de Ciências Biológicas, Universidade de Brasília, Asa Norte, Brasília 70910-900, Brazil. [5] Lab. Biomass Conversion, Embrapa Pantanal, R. 21 de Setembro 1880, 79320-900 Corumbá/MS, Brazil. [6] Cell Biology and Metabolomics Platforms, Gustave Roussy Cancer Campus, Villejuif 94805, France. [7] Equipe 11 labellisée Ligue Nationale contre le Cancer, Centre de Recherche des Cordeliers, Paris 75006, France. [8] Institut National de la Santé et de la Recherche Médicale (INSERM), Paris 75654, France. [9] Université Paris Descartes, Sorbonne Paris Cité, Paris 75015, France. [10] Université Pierre et Marie Curie, Paris 75005, France. [11] Pôle de Biologie, Hôpital Européen Georges Pompidou, AP-HP, Paris 75015, France. [12] Department of Women's and Children's Health, Karolinska University Hospital, Stockholm SE-171 76, Sweden. Jônatas Abrahão, Lorena Silva and Ludmila Santos Silva contributed equally to this work. Correspondence and requests for materials should be addressed to D.R. (email: didier.raoult@gmail.com) or to B.S. (email: bernard.la-scola@univ-amu.fr)

Translation is one of the canonical frontiers between the cell world and the virosphere. Even the simplest non-viral obligatory intracellular parasites present a wealthy set of translation apparatus, including aminoacyl tRNA synthetases, tRNAs, peptide synthesis factors and ribosomal proteins[1–3]. The nature of the parasitism of the most of non-viral obligatory intracellular parasites relies on partial or severe deficiency of genes related to energy production. Although sharing a similar lifestyle with those organisms, the most of the virus lack not only energy production genes, but also translation-related genes[1–3].

In this context, the discovery of mimiviruses and other amoeba-infecting giant viruses surprised the scientific community due their unusual sizes and long genomes, able to encode from hundreds to thousands of genes, including tRNAs, peptide synthesis factors and, for the first time seen in the virosphere, aminoacyl tRNA synthetases (aaRS)[4]. Although the very first discovered mimivirus (*Acanthamoeba polyphaga mimivirus*) encodes four different types of aaRS (Arg, Cys, Met, TyrRS), other mimivirus isolates genomes were described, containing up to seven types of aaRS, as Megavirus chilensis and LBA111 (Arg, Asn, Cys, Ile, Met, Trp, TyrRS)[5]. While amoeba-infecting mimiviruses present up to six copies of tRNA, *Cafeteria roenbergensis virus* (CroV), a group II algae-infecting of *Mimiviridae*, encodes 22 sequences for five different tRNAs[5]. Even more surprisingly, metagenomics data reveal that klosneuvirus genome may encodes aaRS with specificities for 19 different amino acids, over 10 translation factors and several tRNA-modifying enzymes[6]. However, klosneuviruses were not isolated and therefore there is no data regarding important biological features as virus particles, morphogenesis and host-range.

Here we report the discovery of two Tupanvirus strains, the longest tailed *Mimiviridae* members isolated in amoebae. Their genomes are 1.44–1.51 Mb linear double-strand DNA coding for 1276–1425 predicted proteins. Tupanviruses share the same ancestors with mimivirus lineages and these giant viruses present the largest translational apparatus within the known virosphere, with up to 70 tRNA, 20 aaRS, 11 factors for all translation steps, and factors related to tRNA/mRNA maturation and ribosome protein modification. Moreover, two sequences with significant similarity to intronic regions of 18 S rRNA genes are encoded by the tupanviruses and highly expressed. In this translation-associated gene set, only the ribosome is lacking. Tupanvirus is also cytotoxic and causes a severe shutdown of ribosomal RNA and a progressive degradation of the nucleus in host and non-host cells. The analysis of tupanviruses constitutes a new step towards understanding the evolution of giant viruses.

Results

Tupanviruses description and cycle. While attempting to find new and distant relatives of currently known giant viruses, we performed prospecting studies in special environments. Soda lakes (Nhecolândia, Pantanal biome, Brazil) are known as environments that conserve and/or mimic ancient life conditions (extremely high salinity and pH) and are considered some of the most extreme aquatic environments on Earth[7]. We also prospected giant viruses in ocean sediments collected at a depth of 3000 m (Campos dos Goytacazes, Brazilian Atlantic Ocean).

Both in soda lake and deep ocean samples, we found optically visible *Mimiviridae* members that surprisingly harbored a long, thick tail (Fig. 1a, b) as they grew on *Acanthamoeba castellanii* and *Vermamoeba vermiformis*. We named these strains Tupanvirus soda lake and Tupanvirus deep ocean as a tribute to the South American Guarani Indigenous tribes, for whom Tupan—or Tupã—(the God of Thunder) is one of the main mythological figures. Electron microscopy analyses revealed a remarkable virion structure. Tupanviruses present a capsid similar to that of amoebal mimiviruses in size (~450 nm) and structure, including a stargate vertex and fibrils[8] (Figs. 1a–d, 2a–q). However, the Tupanvirus virion presents a large cylindrical tail (~550 nm extension; ~450 nm diameter, including fibrils) attached to the base of the capsid (Figs. 1b–d, 2a,i–k). This tail is the longest described in the virosphere[9]. Microscopic analysis suggests that the capsid and tail are not tightly attached (Figs. 1e, f, 2a, i; Supplementary Movies 1 and 2), although sonication and enzymatic treatment of purified particles were not able to separate the two parts (Fig. 2f–h). The average length of a complete virion is ~1.2 µm, although some particles can reach lengths of up to 2.3 µm because of the variation in the tail's size; this makes them the one of the longest viral particles described to date (Figs. 1, 2i–k). Furthermore, there is a lipid membrane inside the capsid, which is associated with the fusion with the cellular membrane and the release of capsid content (Fig. 1f). Tail content appears to be released after an invagination of the phagosome membrane inside the tail (Fig. 1e). In contrast

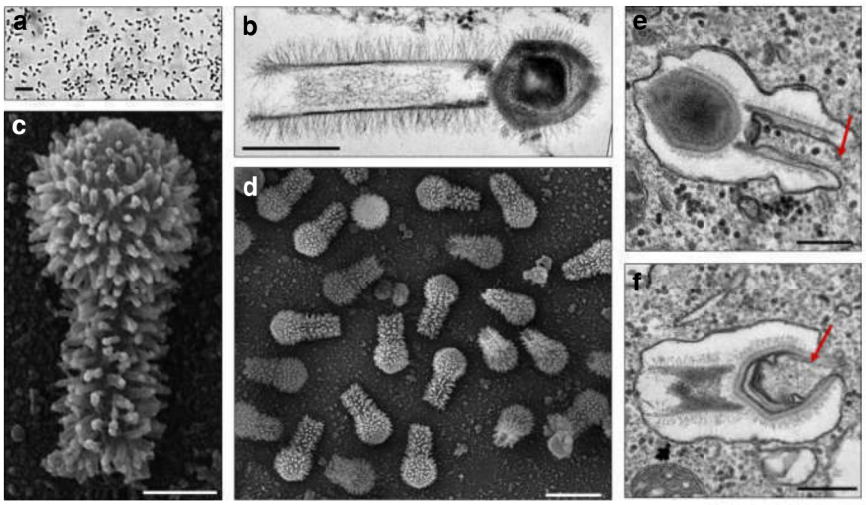


Fig. 1 Tupanvirus soda lake particles and cycle. **a** Optical microscopy of Tupanvirus particles after haemacolour staining (1000 ×). Scale bar, 2 µm. **b** Super particle (>1000 nm) observed by transmission electron microscopy (TEM). Scale bar, 500 nm. **c**, **d** Scanning electron microscopy (SEM) of Tupanvirus particles. Scale bars 250 nm and 1 µm, respectively. **e**, **f** The initial steps of infection in *A. castellanii* involve the release of both capsid (**e**) and tail (**f**) content into the amoeba cytoplasm (red arrows). Scale bars, 350 nm and 450 nm, respectively

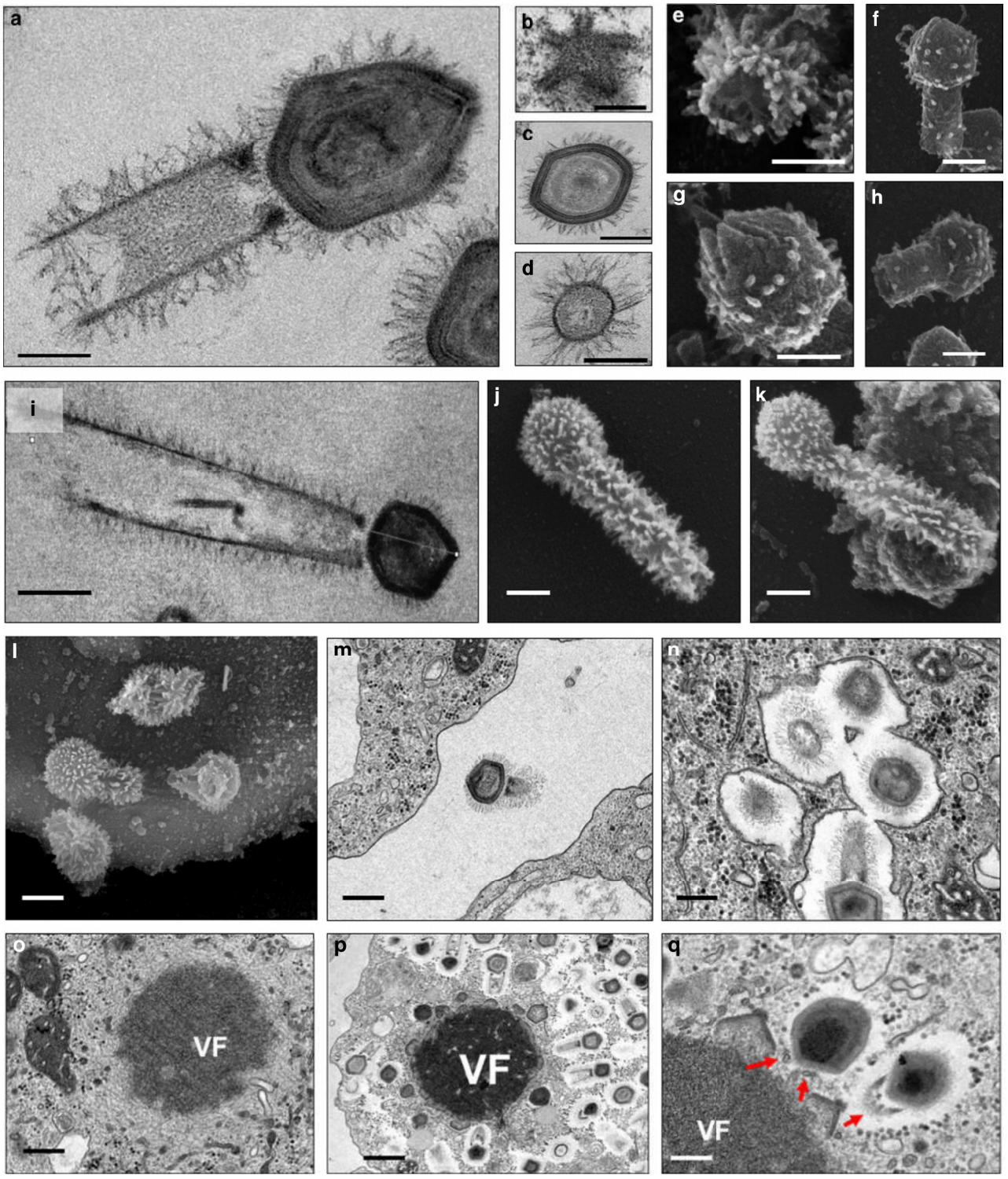


Fig. 2 Tupanvirus soda lake particles and cycle features. **a** Transmission electron microscopy (TEM) highlights the inner elements of the whole particle. Scale bar, 200 nm. **b** Star-gate vertex transversally cut. Scale bar, 100 nm. **c** Capsid transversally cut. Scale bar, 100 nm. **d** Tail transversally cut. Scale bar, 200 nm. **e–h** Scanning electron microscopy (SEM) of purified particles. Scale bars, 250 nm. The treatment of particles with lysozyme, bromelain and proteinase-K removed most of the fibers, revealing head and tail junction details. Super particles (>2000 nm) could be observed by TEM (**i**) and SEM (**j**, **k**). Scale bars, 400 nm. Cycle steps are shown from **l–r**. **l** Viral particle attachment to *Acanthamoeba castellanii* surface; scale bar, 500 nm; **m** phagocytosis; scale bar, 500 nm; **n** particles in a phagosome; scale bar, 500 nm; **o** early viral factory; scale bar, 500 nm; and **p**, **q** mature viral factories. Scale bars 1 µm and 250 nm, respectively. Arrows highlight tail formation associated with the viral factories. VF viral factory

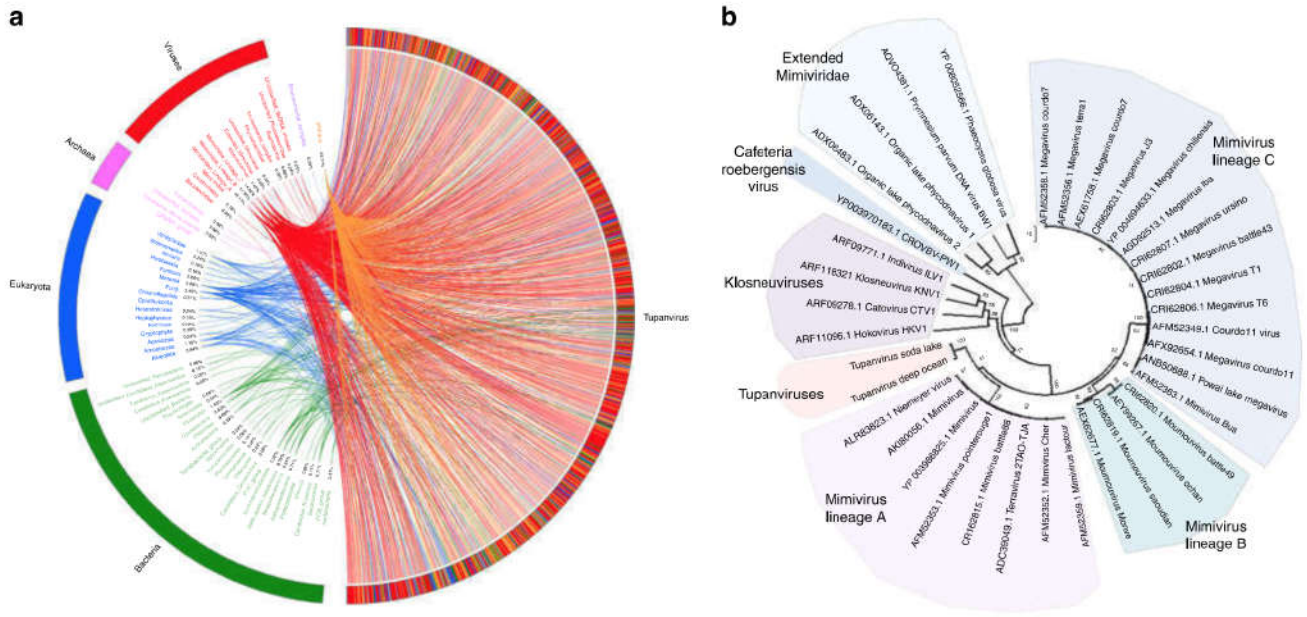


Fig. 3 Tupanvirus soda lake rhizome and Mimiviridae family B DNA polymerase tree. **a** The rhizome shows that most Tupanvirus soda lake genes have mimiviruses as best hits. However, correspondence among Tupanvirus and *Archaea*, *Eukaryota*, *Bacteria* and other viruses was also observed. **b** Family B DNA polymerase maximum likelihood phylogenetic tree demonstrating the position of Tupanvirus among *Mimiviridae* members, likely forming a new genus

to other giant viruses[10–14], Tupanviruses similarly replicate both in *A. castellanii* and *V. vermiformis*. Viral particles attach to the host-cell surface and enter through phagocytosis (1 h.p.i.) (Fig. 2 l,m,n). The inner membrane of the capsid merges with the amoebal host phagosome membrane, releasing the genome (2–6 h.p.i.) (Fig. 1f). A viral factory (VF) is then formed (7–12 h. p.i.)[15] where particle morphogenesis occurs (Fig. 2o–q; Supplementary Movies 1 and 2). During this step, the virion tail is attached to the capsid after its formation and closure. Later in the process (16–24 h.p.i.), the amoebal cytoplasm is filled with viral particles, followed by cell lysis and the release of viruses (Supplementary Fig. 1). This nucleus-like viral factory has also recently been reported in bacteria and fuels the concept of a virocell[16,17]. In that perspective, viral factories actively producing the progeny could be considered as the nuclei of virocells[16,17].

The genomes of tupanviruses. The tupanvirus soda lake and tupanvirus deep ocean genomes (GenBank accession number KY523104 and MF405918) are linear dsDNA molecules measuring 1,439,508 bp and 1,516,267 bp (GC% ~28%), respectively—the fourth largest viral genomes described to date[10,18]—containing a total of 1276 and 1425 predicted ORFs, 375 and 378 of which are ORFans (ORFs with no matches in current databases), respectively. To date, the largest genomes belong to pandoravirus isolates, and the largest one, *P. salinus*, has 2,473,870 bp and encodes 2,556 putative proteins[10]. The rhizome[19] of tupanvirus (graphical representation of gene-by-gene best hits) revealed sequences from mimiviruses of amoebae (~42%) and klosneuviruses (~8%) as their main best hits. Other best hits were mostly sequences from eukaryotes (~11%) and bacteria (~8%) (Fig. 3a; Supplementary Fig. 2A). Tupanviruses exhibited relatively close numbers of best matches to amoebal mimiviruses from lineages A (~10%), B (~18%) and C (~14%). These data and phylogenetic analyses demonstrate that they cluster with amoebal mimiviruses, suggesting that tupanviruses are distant relatives of, and comprise a sister group to, mimiviruses of amoeba (Figs. 3b, 4a). The 'AAAATTGA' promoter motif was found ~410 times in Tupanvirus deep ocean intergenic regions, a frequency similar to that of other *Mimiviridae* members, and ~600% more frequent than those coding regions ($p < 10^{-95}$, Fisher exact test)[20–22] (Fig. 4b). The pangenome of the family *Mimiviridae*, when taking into account the gene contents of tupanviruses, klosneuviruses and distant relatives to amoebal mimiviruses, was found by Proteinortho to comprise 8,753 groups of orthologues ($n = 3588$) or virus unique genes ($n = 5165$). A total of 189 groups were shared by at least one tupanvirus; one mimivirus of *Acanthamoeba* of each lineage A, B and C; and one klosneuvirus, and 100 of them were also shared with Cafeteria roenbergensis virus. In addition, a total of 757 groups were shared by a tupanvirus and at least another mimivirus: 477 were shared with Megavirus chiliensis, 434 with Moumouvirus, 431 with Mimivirus, 287 with Klosneuvirus, 126 with Cafeteria roenbergensis virus, and 59 with Phaeocystis globosa virus 12 T. Among these 757 groups, 583 corresponded to clusters of orthologous groups previously delineated for mimiviruses[19]. Finally, a total of 775 tupanvirus genes were absent from all other mimivirus genomes (Supplementary Data 1).

Proteomic analysis. Proteomic analysis of Tupanvirus soda lake particles revealed 127 proteins, nearly half (67/127 = 52.8%) of which are unknown and eight of which are encoded by ORFans (11/127 = 8.6%) (Supplementary Data 2; Supplementary Note 1). Although 62 Tupanvirus virion proteins homologous were not found by proteomics, either in Mimivirus or in *Cafeteria roenbergensis virus* particles, there are no distinct clues about which protein (s) could be associated with the tupanvirus fibrils and tail structure.

The most complete translational apparatus of the virosphere. Analyses of the tupanvirus gene sets related to energy production revealed a clear dependence of these viruses on host energy production mechanisms, similarly to other mimiviruses, because genes involved in glycolysis, the Krebs cycle and the respiratory chain are mostly lacking[22–24]. Astonishingly, Tupanvirus soda lake and Tupanvirus deep ocean exhibit the largest viral sets of genes involved in translation, with 20 ORFs related to aminoacylation (aaRS) and transport, and 67 and 70 tRNA associated with 46 and 47 codons, respectively (Fig. 5a; Supplementary

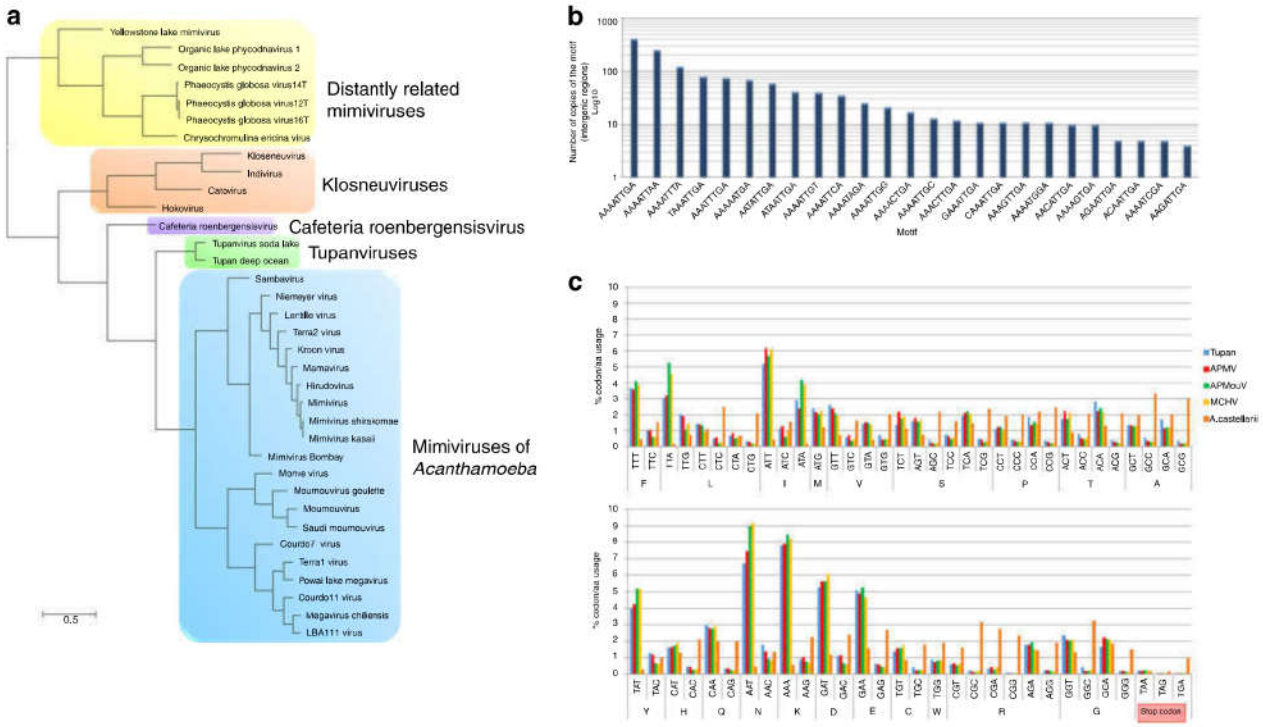


Fig. 4 Tupanvirus soda lake hierarchical clustering, promoter's motifs and amino-acid usage analysis. **a** Hierarchical clustering tree based on presence-absence matrix of cluster of orthologous genes shared by *Mimiviridae* members. **b** Frequency of mimivirus AAAATTGA canonical promoter motifs in Tupanvirus intergenic regions. We also analyzed the presence of the AAAATTGA motif with SNPs, considering each motif position. **c** Comparative amino-acid usage analysis of Tupan, *A. castellanii* and mimivirus lineages **a**, **b** and **c**. The amino-acid usage for protein sequences was calculated using the CGUA (General Codon Usage Analysis) tool

Fig. 2B, C, 3 and 7; Supplementary Datas 3 and 4). Tupanvirus deep ocean encodes an ambiguous tRNA related to the rare amino-acid pyrrolysine. Only selenocysteine-related genes are lacking, as previously observed for many cellular organisms[25]. Several translation factors were identified, including eight translation initiation factors (IF2 alpha, IF2 beta, IF2 gamma, IF4e (2 copies in Tupanvirus soda lake), IF5a (2 copies in Tupanvirus deep ocean), SUI1, IF4a), one elongation/initiation factor (GTP-binding elongation/initiation), one elongation factor (Ef-aef-2), and one release factor (ERF1) (Fig. 5a; Supplementary Fig 2B, C and 3; Supplementary Datas 3 and 4). Furthermore, we detected additional translation-related genes: factors related to tRNA maturation and stabilization (tRNA nucleotidyltransferase, tRNA guanylyltransferase, cytidine deaminase, RNA methyl transferase); mRNA maturation (poly(A) polymerase, mRNA capping enzyme) and splicing (RNA 2—phosphotransferase Tpt1 family protein); and ribosomal protein modification (ribosomal-protein-alanine N-acetyltransferase, FtsJ-like methyl transferase) (Fig. 5a; Supplementary Fig 2B; Supplementary Datas 3, 4). Phylogenies based on aaRS are acknowledged to be complex as they depend on sequence sampling, and natural and canonical taxonomical groups can be separated into different clades, as previously observed[6]. To reduce such disturbances in the trees during their construction, we selected the 100 best hits related to Tupanvirus soda lake, plus the sequences of 5–10 amoebozoa and those of klosneuviruses. In most of the trees, sequences from natural clades were clustered together, although some amoebozoa were separated into different clusters. Based on the 20 aaRS trees, it is not possible to state that the origin of most of these tupanvirus aaRS genes is cellular (Supplementary Fig. 4). Furthermore, the mosaic structure of aaRS gene reinforced the difficult to state on the origin of these genes, as illustrated in Supplementary Fig. 5. This contrasting result to that reported by Schultz et al.[6], 2017 for klosneuviruses may also be explained by the different sampling used for alignments and trees construction. In addition, the codon and amino-acid usage of tupanviruses is substantially different from that of *Acanthamoeba* spp. In tupanviruses, we observed a high correlation between tRNA isoacceptors and the most used codons, as these viruses present more tRNA isoacceptors for highly abundant codons (Fig. 4c; Supplementary Figs. 2C and 3). Surprisingly, we found two different copies of an 18 S rRNA intronic region in tupanviruses (Supplementary Fig. 6; Fig. 6a–e). In fact, such 18 S rRNA intronic regions are widespread in all mimiviruses (lineages A and B present only one copy in an intronic region and next to a self-splicing group I intron endonuclease) but are not found in klosneuviruses. Phylogenetic analyses revealed that the two copies found in tupanviruses had separate and different origins (Fig. 6f). Although Tupanvirus 18 S rRNA intronic sequences are located in intergenic regions, qPCR and FISH demonstrated that Tupanvirus soda lake 18 S rRNA intronic sequences are highly expressed during the entire replicative cycle but particularly during intermediate and late phases (6 and 12 h post-infection) (Fig. 7a; Supplementary Fig. 8). Furthermore, Tupanvirus soda lake is more tolerant to the translation-inhibiting drugs geneticin and cycloheximide than Mimivirus, an impressive characteristic considering the natural ribosomal RNA shutdown it performs in the permissive host (Fig. 8h). The functions of these 18 S rRNA intronic region sequences require further clarification. No exonic region of 18 S rRNA was found in the genomes of tupanviruses or any previously described mimivirus. The comparison between contents in translation-related categories of genes present in tupanviruses and cellular organisms reveals that tupanviruses present a richer gene set than *Candidatus Carsonella ruddii* (*Bacteria*) and *Nanoarchaeum equitans* (*Archaea*) (not considering ribosomal proteins). Tupanvirus deep ocean has even

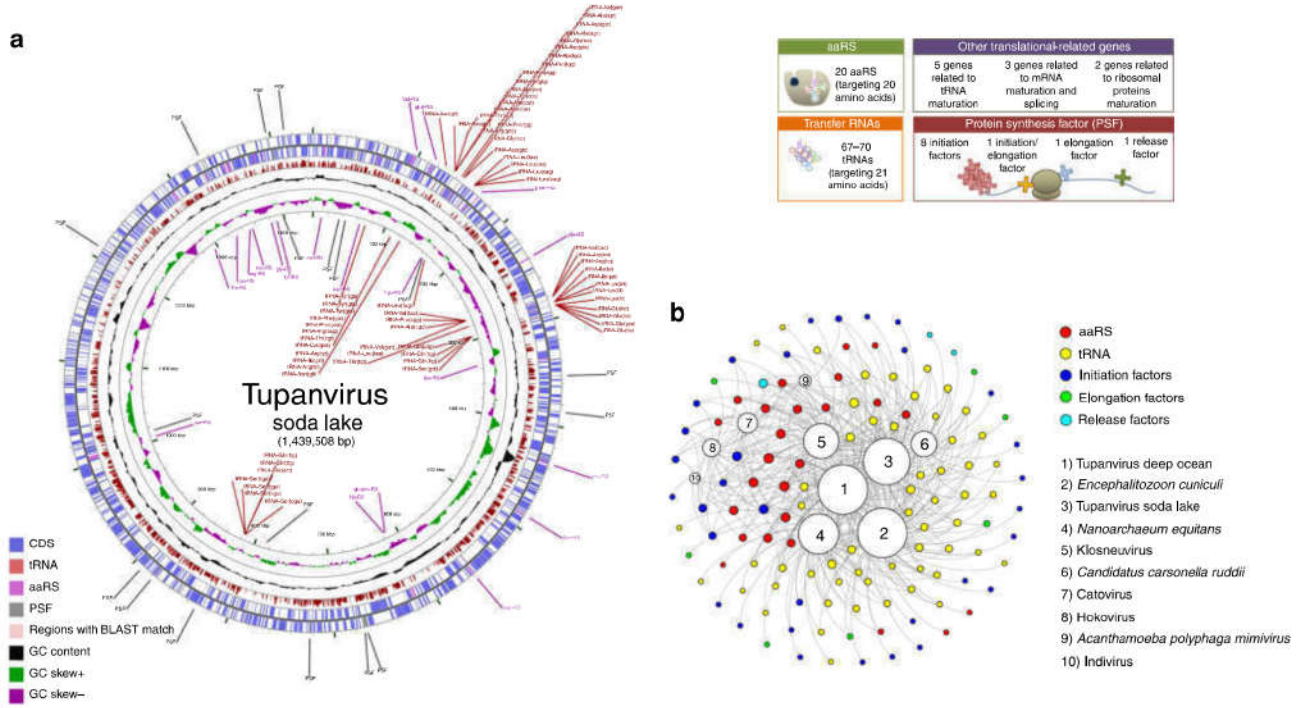


Fig. 5 Tupanvirus genome-translation-related factors. **a** Circular representation of Tupanvirus soda lake genome highlighting its translation-related factors (aaRS, tRNAs and PSF). The box (upright) summarizes this information and considers the Tupanvirus deep ocean data set. **b** Network of shared categories of translation-related genes (not considering ribosomal proteins) present in tuparviruses, Mimivirus (APMV), Klosneuvirus, Catovirus, Hokovirus, Indivirus and cellular world organism—*Encephalitozoon cuniculi* (*Eukaryota*), *Nanoarchaeum equitans* (*Archaea*) and *Candidatus Carsonella ruddii* (*Bacteria*). The diameter of the organism's circles (numbers) is proportional to the number of translatior-related genes present in those genomes. CDS coding sequences, tRNA transfer RNA, aaRS aminoacyl tRNA synthetase, PSF protein synthesis factors

more such genes than *Encephalitozoon cuniculi*, a eukaryotic organism (Fig. 5b). Even if the impressive translation gene set recently described for klosneuviruses is taken into account, tupanviruses are the first viruses reported to harbor the complete set of the 20 aaRS (Fig. 5; Supplementary Figs. 3, 6 and 7).

Host range and host-ribosomal shutdown characterization. In contrast to other mimiviruses, Tupanvirus soda lake was able to infect a broad range of protist organisms in vitro. Surprisingly, we observed four distinct profiles of infectiveness: (i) productive cycle in permissive cells; (ii) abortive cycle; (iii) refractory cells; and (iv) most surprisingly, non-host cells exhibiting a cytotoxic phenotype in the presence of Tupanvirus without multiplication, a circumstance never previously reported—to the best of our knowledge (Supplementary table 1; Supplementary Notes 1). This latter profile was intriguing because toxicity was observed in *Tetrahymena* sp., a ravenous free-living protist[26]. This unusual phenotype was also observed in *A. castellanii* but only at higher multiplicities of infection (50 and 100; Fig. 8a). No such cytotoxicity was observed for Mimivirus (Fig. 8a). This toxic profile associated specifically with Tupanvirus appears to be related to shutting down host ribosomal RNA abundance, insofar that Tupanvirus leads to a reduction of host rRNA amounts by a mechanism not related to the canonical ribophagy/autophagy process (Figs. 7b–d, 8b–d; Supplementary Fig. 9). A remarkable acidification of amoebal cytoplasm is induced by Tupanvirus infection (but not mimivirus) (Fig. 7b; Supplementary Fig. 9A, B). The treatment of acanthamoeba with chloroquine, a lysosomal toxin, or bafilomycin did not prevent the Tupanvirus-induced rRNA shutdown (Supplementary Fig. 9A, C–H). Transfection with an siRNA targeting Atg8-2, a protein required for ribophagy/autophagy, failed to prevent Tupanvirus-induced ribosomal shutdown in *A. castellanii* (Fig. 7c, d). Transmission electron microscopy (TEM) of vesicles containing ribosomes after Tupanvirus infection (2 h.p.i.) revealed that these structures were formed close to the nuclear membrane after invagination and engulfment of ribosomes, mostly by single-membrane vesicles, rarely by double-membrane vesicles (Fig. 7e). These small vesicles then aggregated, accumulating more ribosomes and forming large structures containing ribosomes (Fig. 8d), which were fully depleted from the amoebal cytoplasm 6–9 h.p.i, when strong cellular rRNA shutdown could be detected (Fig. 8c). In addition, Tupanvirus infection induces nuclear/nucleolar progressive degradation, which is temporally associated with cellular rRNA shutdown (Fig. 7f). Mimivirus infection also caused changes in nucleolus architecture, but such alterations were not comparable to those observed upon Tupanvirus infection (Fig. 7e, f). Although the formation of vesicles containing ribosomes and nuclear degradation can be related to cellular rRNA shutdown during Tupanvirus infection, the mechanisms involved in cellular rRNA degradation after the formation of large vesicles containing ribosomes remain to be investigated. One possibility that should be explored is that Tupanvirus might preferentially target some ribosomes to favor the translation of its own (as opposed to cellular) proteins, as previously described for poxviruses[27].

The toxicity effect and rRNA shutdown are independent of Tupanvirus replication; instead, they are caused by the viral particle (Fig. 8e–g). UV-light-inactivated particles continued to induce the depletion of *Acanthamoeba* rRNA (Fig. 8g). Although Tupanvirus is not able to replicate within *Tetrahymena* sp., it is phagocytosed in a voracious manner (as are other available macromolecular structures) and forms large intracellular vesicles, where the capsid and tail release their content inside the protist cytoplasm (Fig. 8i–k). The virus induces gradual vacuolization (Fig. 8i), loss of motility, a decrease in the phagocytosis rate

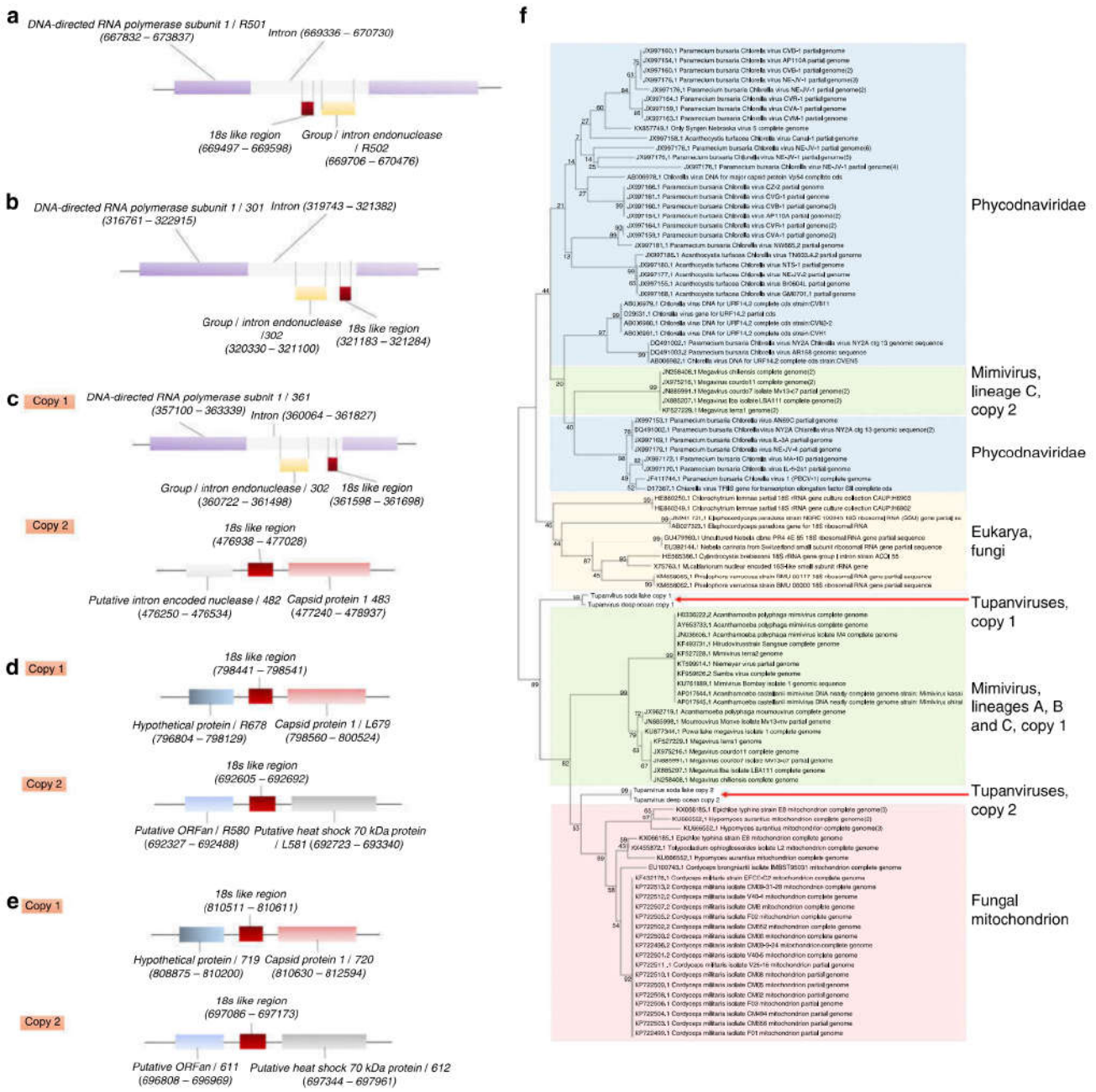


Fig. 6 Adjacent regions of 18S rRNA intronic sequences in the genus *Mimivirus* and Tupanvirus and maximum likelihood phylogenetic tree of 18S rRNA intronic region. Core sequences are represented for lineages A (**a**), B (**b**), C (**c**), Tupanvirus soda lake (**d**) and Tupanvirus deep ocean (**e**). Phylogenetic tree of 18S rRNA intronic region present in mimivirus (**e**), *Phycodnaviridae*, eukaryotes and fungi mitochondrion

(Fig. 8m), a decrease in rRNA abundance (Fig. 8l) and triggers nuclear degradation (Fig. 7g), similar to the effects observed in *A. castellanii* cells with high multiplicities of infection (Fig. 8a–d). We performed in vitro simulations to determine whether this toxicity could affect the maintenance of Tupanvirus populations in a solution containing both *Tetrahymena* and *Acanthamoeba*. Our data suggest that at M.O.I. of 10 the reduction of the physiological activity of *Tetrahymena* sp., a (non-host) predator, decreases the ingestion of Tupanvirus particles (Fig. 8n), improving their chances to find its host, *Acanthamoeba*, and keeping viral titers constant in the system. In contrast, mimivirus particles, which do not present any toxicity to *Tetrahymena* sp., are quickly predated, and the virus is eliminated from the system after some days (Fig. 8n). Although we have no clues about the hosts of tupanviruses at their original habitats nor if such high M. O.I. would be expected in nature, to our knowledge, a viral particle responsible for the modulation of host and non-host organisms independent of viral replication has not been previously described.

Discussion

Considering that tupanviruses comprise a sister group to amoebal mimiviruses, we can hypothesize that the ancestors of these clades of *Mimiviridae* could had a more generalist lifestyle and were able to infect a wide variety of hosts. In this view, the ancestors of tupanviruses (and maybe of amoebal mimiviruses) might have already been giant viruses that underwent reductive evolution,

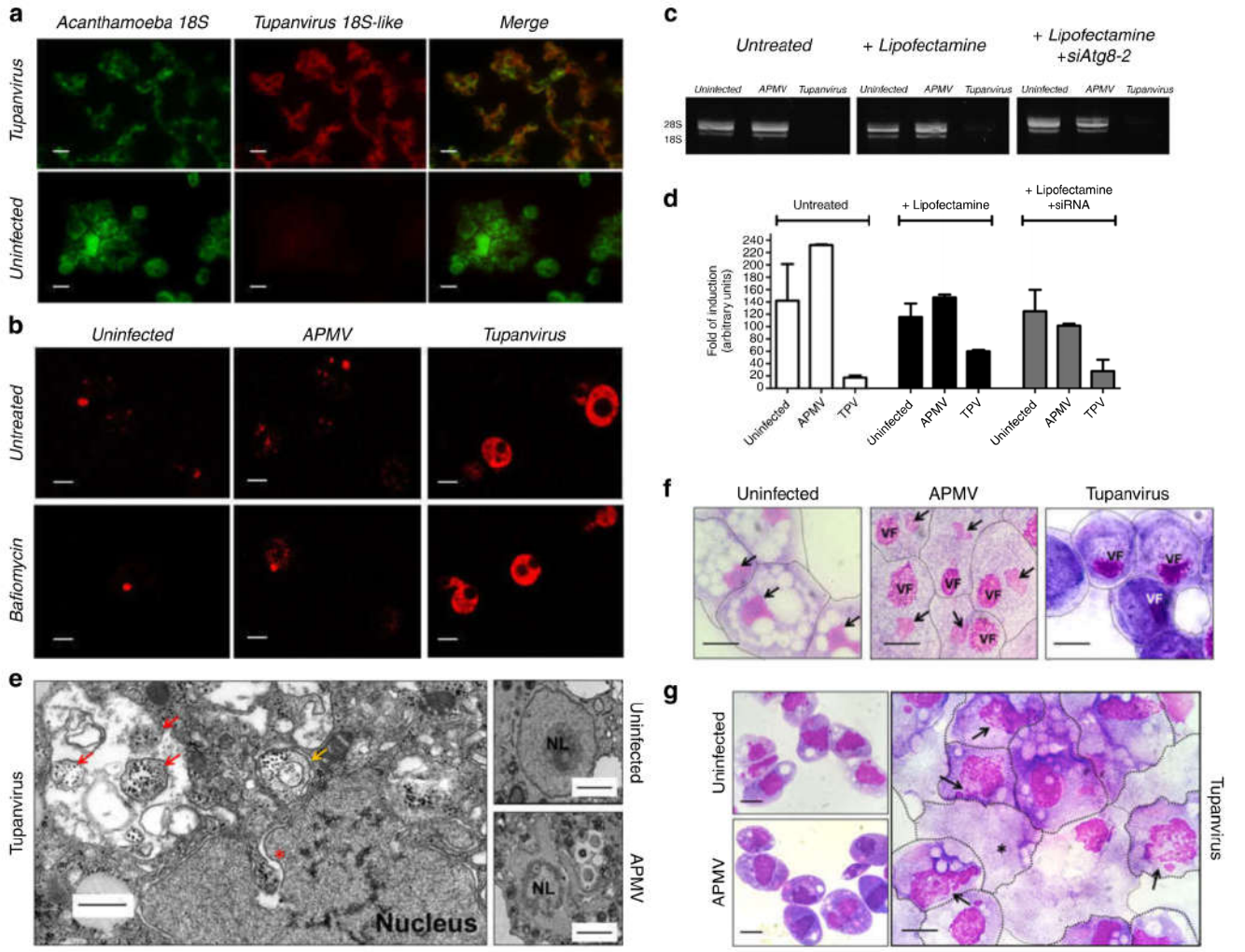


Fig. 7 Tupanvirus soda lake biological features in a host (*A. castellanii*) and non-host (*Tetrahymena* sp.). **a** Expression of Tupanvirus 18S intronic sequence-copy 1 transcript, 12 h post-infection observed by fluorescence in situ hybridization (FISH) (red). Tupanvirus-induced shutdown of *A. castellanii* ribosomal RNA18S transcripts (green). Scale bars, 10 µm. **b** Tupanvirus, but not mimivirus, induces strong acidification of *A. castellanii* cytoplasm (9 h.p.i.), even in the presence of bafilomycin A1. Scale bars, 10 µm. **c**, **d** The silencing of the canonical autophagy protein Atg8-2 does not prevent rRNA shutdown induced by Tupanvirus infection. Error bars, standard deviation. **e** Electron microscopy of *A. castellanii* infected by Tupanvirus (2 h.p.i.), highlighting the degradation of the nucleolus, nuclear disorganization and the formation of ribosome-containing vesicles near nuclear membranes. Scale bar, 500 nm. Red arrow: single-membrane vesicles containing ribosomes; orange arrow: double-membraned vesicles containing ribosomes; asterisk: ribosomes wrapping by the external nuclear membrane. Right: electron microscopy of uninfected amoebae and APMV infected cell (8 h.p.i.) showing a mild nucleolar disorganization in the presence of the virus. Scale bars: 1 µm. **f** Haemacolour staining showing the nuclear degradation of *A. castellanii* induced by Tupanvirus infection (9 h.p.i.) compared with infection by mimivirus (APMV) and uninfected cells. Tupanvirus-infected cells are purplish because of cytoplasm acidification. Scale bars, 5 µm. Arrows: nucleus, when present; VF: viral factory. **g** Haemacolour staining showing the nuclear degradation in *Tetrahymena sp.* induced by Tupanvirus infection (4 days post-inoculation) compared with mimivirus (APMV)-inoculated cells and uninfected cells. Tupanvirus-infected cells present an atypical shape and intense vacuolization, and some cells lack a nucleus (asterisk). Arrows: nucleus under degradation. Scale bars, 5 µm. The experiments were performed 3 times independently, with two replicates each

although some genes could have been acquired over time, as previously hypothesized for other mimiviruses[5,23,28,29]. A reductive evolution pattern is typical among obligatory intracellular parasites[30–32]. In these cases, the organisms lose genes related to energy production, which is one of the main reasons for their obligatory parasitic lifestyle. In an alternative scenario, a simpler ancestor could had substantially acquired genes over time and became more resourceful, being able to infect a broader host range[5,33]. Nevertheless, tupanvirus presents the most complete translational apparatus among viruses, and its discovery takes us one step forward in understanding the evolutionary history of giant viruses.

Methods

Virus isolation and host-range determination. In April 2014, a total of 12 sediment samples were collected from soda lakes in Southern Nhecolândia, Pantanal, Brazil. In 2012, 8 ocean sediments samples were collected from 3000 meters below water line surface at Bacia de Campos, in Campos dos Goytacazes, Rio de Janeiro, Brazil. The collection was performed by a submarine robot, during petroleum prospection studies performed by the Petrobras Company, and kindly provided to our group. The samples were stored at 4 °C until the inoculation process. The samples were transferred to 15 mL flasks, and 5 mL of Page's Amoebae Saline (PAS) was added. The solution was stored for 24 h to decant the sediment. The liquid was then subjected to a series of filtrations: first through paper filter and then through a 5 µm filter to remove large particles of sediment and to concentrate any giant viruses present. For co-culture, the cells used were *A. castellanii* (strain NEFF) and *V. vermiformis* (strain CDC 19), purchased from ATCC. These cell strains

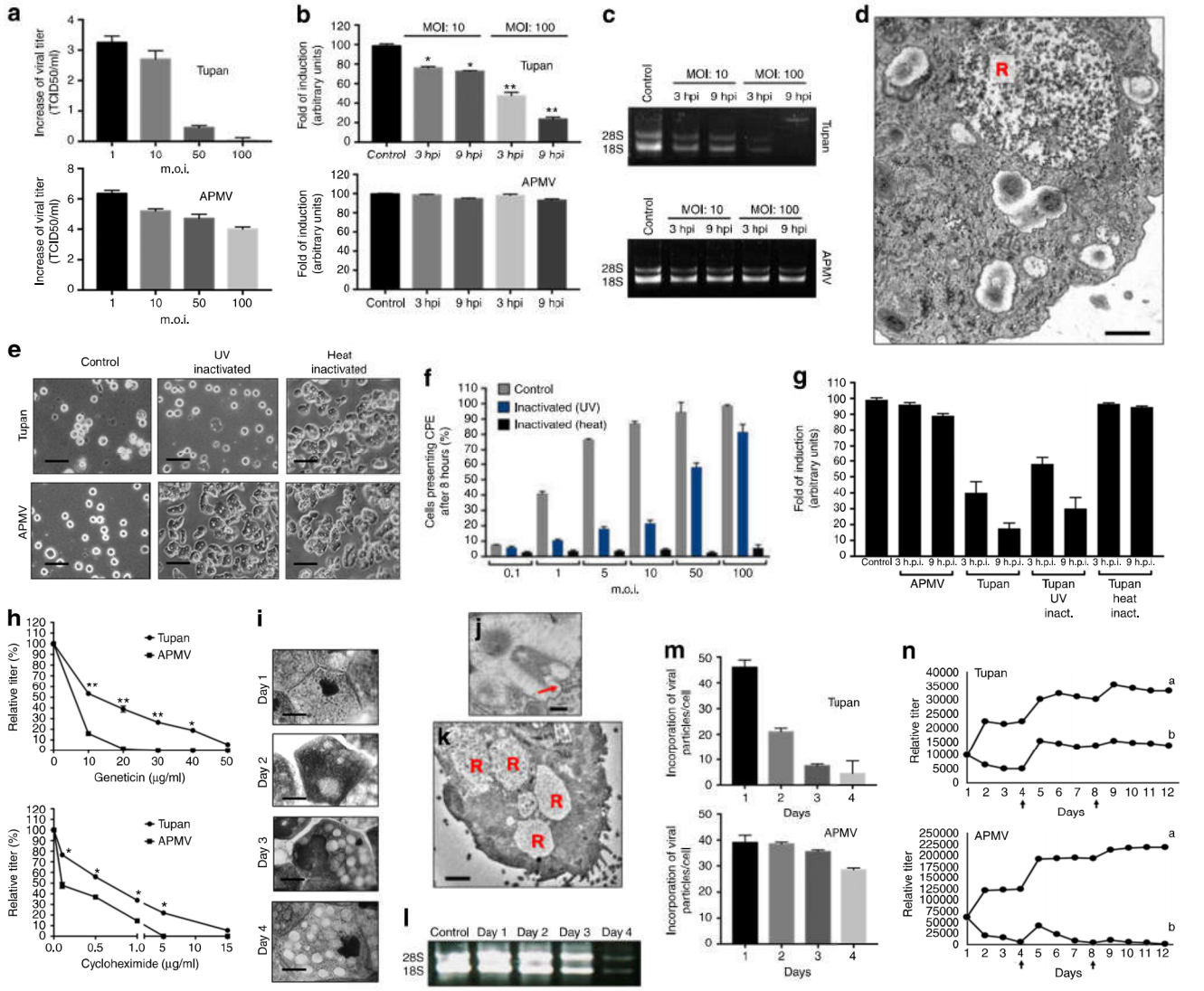


Fig. 8 rRNA shutdown induced by Tupanvirus and toxicity assays. **a** Tupanvirus and mimivirus titers (log_{10} values) 24 h.p.i. in *Acanthamoeba castellanii* at distinct MOIs. **b** Ribosomal 18S RNA relative measure by qPCR from *A. castellanii* infected by tupanvirus or mimivirus at an MOI of 10 or 100, 3 and 9 h post-infection. **c** Electrophoresis gel showing ribosomal 18S and 28S RNA from *A. castellanii* under the same conditions described in **b**. **d** Vesicle containing a large amount of *A. castellanii* ribosomes (R) 6 h after Tupanvirus infection. Scale bar, 1 μm. (**e**) Cytopathic effect of *A. castellanii* inoculated with Tupanvirus or mimivirus after UV or heat inactivation, MOI of 100, 8 h post-inoculation. Scale bar, 20 μm. **f** Counting of *A. castellanii* presenting cytopathic effect 8 h post-inoculation with tupanvirus inactivated by UV or heating under different MOIs. **g** Ribosomal 18S RNA relative measure by qPCR from *A. castellanii* infected by UV-inativated or heat-inactivated Tupanvirus, or APMV, at an MOI of 100, 3 and 9 h post-infection. **h** Dose-response of Tupanvirus and APMV in *A. castellanii* pre-treated with distinct doses of geneticin or cycloheximide. (**i**) Progressive vacuolization of tetrahymena after infection with Tupanvirus (days 1–4). Scale bars, 10 μm. **j** Tupanvirus tail content releasing in tetrahymena 1 h post-inoculation (TEM). Scale bar, 200 nm. **k** Vesicles containing large amounts of *Tetrahymena* ribosomes (R) after Tupanvirus infection. Scale bar, 3.5 μm. **l** Electrophoresis gel showing rRNA shutdown in tetrahymena inoculated with Tupanvirus at an MOI of 10 (days 1–4). **m** Rate of particles incorporation per tetrahymena cell (days 1–4). **n** Simulations showing the decrease in APMV and maintenance of Tupanvirus populations over the analyzed days after infection of a mix of *Acanthamoeba castellanii* and tetrahymena at an MOI of 10 (lines indicated by 'b' in both graphs of **n**). At days 4 and 8, fresh PYG medium and 10^5 *A. castellanii* were added to the systems (arrows). For the negative control of this experiment we pre-treated tetrahymena with 20 μg/ml of geneticin (lines indicated by 'a' in both graphs of **n**). In this case, both APMV and Tupanvirus were able to grow in the system. Statistical analyses in **b** and **h**: t-test based on control groups (**b**) or corresponding virus/drug concentrations (**h**). *:$p < 0.05$; **:$p < 0.01$. The experiments were performed 3 times independently, with two replicates each. Error bars (**a**, **b**, **f**, **g**, **h**, **m** and **n**), standard deviation

were stored in 75 cm^2 cell culture flasks containing 30 mL of peptone yeast extract glucose medium (PYG) at 28 °C. After 24 h of growth, cells were harvested and pelleted by centrifugation. The supernatant was removed, and the amoebae were resuspended three times in sterile PAS. After the third washing, 500,000 *A. castellanii* or *V. vermiformis* were resuspended in PAS or TS solutions and seeded in 24-well plates. The amoebae suspensions were added to an antibiotic mix containing ciprofloxacin (20 μg/mL; Panpharma, Z.I., Clairay, France), vancomycin (10 μg/mL; Mylan, Saint-Priest, France), imipenem (10 μg/mL; Mylan, Saint-Priest, France), doxycycline (20 μg/mL; Mylan, Saint-Priest, France), and voriconazole (20 μg/mL; Mylan, Saint-Priest, France). Each 100 μL of sample was mixed and inoculated in the numbered (1–12) wells and incubated at 30 °C in a humid chamber. A negative control was used in each plate. The wells were observed daily under an optical microscope. After 3 days, new passages of the inoculated wells were performed in the same manner until the third passage. In this passage, the content of the wells presenting lysis and cytopathic effects were collected and stored for production and analysis of the possible isolates by haemacolour staining and electron microscopy using the negative stain technique. Of the twelve tested samples from Pantanal, we found Tupanvirus (soda lake) in three. In only one

ocean sample we did isolate Tupanvirus (deep ocean). To evaluate the Tupanvirus soda lake host range, a panel of cell lines was subjected to virus infection at an multiplicity of infection (MOI) of 5: *Acanthamoeba castellanii* (ATCC30010), *Acanthamoeba royreba* (ATCC30884), *Acanthamoeba griffin* (ATCC50702), *Acanthamoeba* sp. E4 (IHU isolate), *Acanthamoeba* sp. Micheline (IHU isolate), *Vermamoeba vermiformis* (ATCC50237), *Dictyostelium discoideum* (ATCC44841), *Willartia magna* (ATCC50035), *Tetrahymena hyperangularis* (ATCC 50254), *Trichomonas tenax* (ATCC 30207), RAW264.7 (Mouse leukemic monocyte-macrophage) (ATCCTIB71) and THP-1 (human monocytic cell line) (ATCCTIB202). Cell lines were tested for mycoplasma. The assays were carried out in 24-well plates, and cells were incubated for 24 or 48 h. The tupanvirus titer was measured in *A. castellanii* by end-point and calculated by the Reed-Muench method[34]. In parallel, the samples were subjected to qPCR, targeting the tupanvirus tyrosyl RNA synthetase (5′-CGCAATGTGTGGAGCCTTTC-3′ and 5′-CCAAGA-GATCCGGCGTAGTC-3′) and aiming to verify viral genome replication (Biorad, CA, USA). Tupanvirus was propagated in 20 *A. castellanii* 175 cm^2 cell culture flasks in 50 mL PYG medium. The particles were purified by centrifugation through a sucrose cushion (50%), suspended in PAS and stored at −80 °C. Purified particles were used for genome sequencing, proteomic analysis[10], and microscopic and biological assays.

Cycle and virion characterization. All biological tests were performed with Tupanvirus soda lake only. To investigate the viral replication cycle by TEM, 25 cm^2 cell culture flasks were filled with 10×10^6 *A. castellanii* per flask, infected by Tupanvirus at a multiplicity of infection of 10 and incubated at 30 °C for 0, 2, 4, 6, 8, 12, 15, 18, and 24 h. One hour after virus-cell incubation, the amoeba monolayer was washed three times with PAS buffer to eliminate non-internalized viruses. A total of 10 mL of the infected cultures was distributed into new culture flasks. A culture flask containing only amoeba was used as the negative control. The infected cells and control sample were fixed and prepared for electron microscopy[14]. For immunofluorescence, *A. castellanii* cells were grown, infected by Tupanvirus at a multiplicity of infection of 1 as described and added to coverslips for 0, 2, 4, 6, 8, 12, 15, 18, and 24 h. After infection, the cells were rinsed in cold phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde (PFA) in PAS for 10 mins. After fixation, cells were permeabilized with 0.2% Triton X-100 in 3% bovine serum albumin (BSA)–PAS for 5 min, followed by rinsing with 3% BSA–PBS three times. Cells were then stained for 1 h at room temperature with an anti-tupanvirus antibody produced in mice (According Aix Marseille University ethics committee rules). After incubation with an anti-mouse secondary antibody, fluorescently labeled cells were viewed using a Leica DMI600b microscope. For Tupanvirus virion characterization, we also used scanning electron microscopy (SEM)[35]. Chemical treatment with proteases and sonication was performed as described elsewhere[35] to investigate fiber composition and the attachment between capsid and tail. For tomography videos, tilt series were acquired on a Tecnai G2 transmission electron microscope (FEI) operated at 200 keV and equipped with a 4096 × 4096 pixel resolution Eagle camera (FEI) and Explore 3D (FEI) software. The tilt range was 100°, scanned in 1° increments. The magnification ranged between 6,500 and 25,000, corresponding to pixel sizes between 1.64 and 0.45 nm, respectively. The image size was 4,096 × 4,096 pixels. The average thickness of the obtained tomograms was 298 ± 131 nm ($n = 11$). The tilt series were aligned using ETomo from the IMOD software package (University of Colorado, USA) by cross-correlation (http://bio3d.colorado.edu/imod/). The tomograms were reconstructed using the weighted-back projection algorithm in ETomo from IMOD. ImageJ software was used for image processing.

Genomes sequencing and analyses. The tupanviruses genomes were sequenced using an Illumina MiSeq instrument (Illumina Inc., San Diego, CA, USA) with the paired end application. The sequence reads were assembled de novo using ABYSS software and SPADES, and the resulting contigs were ordered by the Python-based CONTIGuator.py software. The obtained draft genomes were mapped back to verify the read assembly and close gaps. The best assembled genome was retained, and the few remaining gaps (three) were closed by Sanger sequencing. The gene predictions were performed using the RAST (Rapid Annotation using Subsystem Technology) and GeneMarkS tools. Transfer RNA (tRNA) sequences were identified using the ARAGORN tool. The functional annotations were inferred by BLAST searches against the GenBank NCBI non-redundant protein sequence database (nr) (*e*-value < 1×10^{-3}) and by searching specialized databases through the Blast2GO platform. Finally, the genome annotation was manually revised and curated. The predicted ORFs that were smaller than 50 amino acids and had no hits in any database were ruled out. Tupanvirus codon and aa usages were compared with those of *A. castellanii* and other lineages of mimiviruses. Sequences were obtained from NCBI GenBank and subjected to CGUA (General Codon Usage Analysis). The global distribution of Tupanvirus tRNAs was analyzed and compared manually with viral aa usage considering the corresponding canonical codons related to each aa. Phylogenetic analyses were carried out based on the separate alignments of several genes, including family B DNA polymerase, 18 S rRNA intronic regions (copies 1/2) and 20 aminoacyl-tRNA synthetases (aaRS). The predicted aa sequences were obtained from NCBI GenBank and aligned using Clustal W in the Mega 7.0 software program. Trees were constructed using the maximum likelihood evolution method and 1000 replicates. The analysis of aaRS domains was carried out using NCBI Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). A search for promoter sequences was performed in intergenic regions based on a search for the mimivirus canonical AAAATTGA promoter sequence, as previously described[20]. Single-nucleotide polymorphisms (SNP) in the AAAATTGA promoter sequence were also considered for each base, considering all possibilities. Gene sets available for members of the family *Mimiviridae* and those of Tupanvirus soda lake and Tupanvirus deep ocean were used for analyses of the mimivirus pangenome. Groups of orthologues were determined using the Proteinortho tool V51 with 1*e*−3 and 50% as the *e*-value and coverage thresholds, respectively. Concurrently, BLAST searches were performed using ORFs of all mimivirus genomes available in the NCBI GenBank nucleotide sequence database against the set of clusters of orthologous groups previously delineated for mimiviruses (mimiCOGs) ($n = 898$), with 1*e*−3 and 50% as the *e*-value and coverage thresholds, respectively. For rhizome preparations, all coding sequences were blasted against the NR database, and results were filtered to retain the best hits. Taxonomic affiliation was retrieved from NCBI. For the construction of a translation-associated elements network, the different classes of translation elements of each organism included in the analysis were obtained by searching for each component within their genome, according to protein function annotation using Blastp searches against the GenBank NCBI non-redundant protein sequence database. The tRNA components were obtained using the ARAGORN tool. Different tRNA molecules were included in the analysis, considering the anti-codon sequence. Repeated elements were eliminated to avoid analysis of duplicate events. The layout of the network was generated by a force-directed algorithm—followed by local rearrangement for visual clarity, leaving the network's overall layout unperturbed—using the program Gephi (https://gephi.org).

Ribosomal RNA shutdown and toxicity assays. To investigate the toxicity of Tupanvirus particles, 1 million *A. castellanii* cells were infected with Tupanvirus or mimivirus at a multiplicity of infection of 1, 10, 50, or 100 and incubated at 32 °C. At 0 and 24 h post-infection, the cell suspensions were collected and titered as previously described. A fraction of this suspension (200 µL) was subjected to RNA extraction (Qiagen RNA extraction Kit, Hilden, Germany). The RNA was subjected to reverse transcription by using Vilo enzymes (Invitrogen, CA, USA) and then used as a template in qPCR targeting *A. castellanii* 18 S rRNA (5′-TCCAATTTTCTGCCACCGAA-3′ and 5′-ATCATTACCCTAGTCCTCGCG C-3′). The values were expressed as arbitrary units (delta-Ct). Normalized amounts of the original RNA extracted from each sample were electrophoresed in 1% agarose gel with TBE buffer and run at 150 V. TEM over the entire testing period was performed to evaluate the presence of ribosome-containing vesicles and other cytological alterations. To investigate the nature of virion toxicity, purified Tupanvirus was inactivated by UV light (1 h of exposure, 60 W/m^2) or heating (80 °C, 1 h)—inactivation was confirmed by inoculation on *Acanthamoeba castellanii*, CPE was observed for 5 days and lack of replication was confirmed by qPCR—and inoculated onto *A. castellanii* containing 500,000 cells at multiplicities of infection of 0.1, 1, 5, 10, 50, and 100. The assays were performed in PAS solution. The cytopathic effect was documented and quantified in a counting-cells chamber. Inactivated mimivirus was used for comparison. To determine whether Tupanvirus-induced shutdown of amoebal 18 S rRNA even after inactivation, 500,000 cells were infected (at a multiplicity of infection of 100) and collected at 3 and 9 h post-infection, and amoebal 18 S rRNA levels were measured by qPCR. APMV was used as the control. The sensitivity of Tupanvirus and mimivirus to the translation-inhibiting drugs geneticin and cycloheximide was tested. A total of 500,000 *A. castellanii* cells were pre-treated with different concentrations of the drugs (0–50 and 0–15 µg/ml, respectively) for 8 h and then infected at a multiplicity of infection of 10. Twenty-four hour post-infection, cells were collected, and the viral titers were measured. To investigate the toxicity effect of Tupanvirus particles in the non-host *Tetrahymena* sp., 1 million fresh cells were infected at a multiplicity of infection of 10 in a medium composed of 50% PYG and 50% PAS. The cytopathic effect was monitored for 4 days post-infection, given the reduction of cell movement and vacuolization (lysis or viral replication was not observed). Each day post-infection, 100 µL of infected cell suspension was collected and subjected to cytospin and haemacolour staining to observe vacuolization and other cytological alterations induced by the virus. Other 100 µL aliquots were used to investigate the occurrence of rRNA shutdown induced by Tupanvirus. To this end, the samples were subjected to RNA extraction and electrophoresis. Viral infection in *Tetrahymena* sp. was also observed by TEM at a multiplicity of infection of 10. To determine whether Tupanvirus particles affect the rate of *Tetrahymena* sp. phagocytosis because of toxicity, the rate of viral particle incorporation per cell was calculated during the period of infection. The ratio of TCID50 (infectious entities) and total particles was first calculated by counting the number of viral particles in a counting chamber (approximately 1 $TCID_{50}$ to 63 total particles). One million *Tetrahymena* cells were infected by Tupanvirus or mimivirus at an MOI of 10 $TCID_{50}$. Twelve hour post-infection, the number of viral particles in the medium was estimated by counting the remaining (non-phagocytized) particles. An input of 10 $TCID_{50}$ per cell was added each day post-infection (in separate flasks, one for each day), and the rate of particles phagocytosis was calculated 12 h post-input. For the calculation, the remaining particles from the day before were considered (counted immediately before the input). Considering the toxicity caused by Tupanvirus, but not APMV, in tetrahymena, we conducted an in vitro experiment

aiming to investigate the ability to maintain Tupanvirus or APMV in a system containing both *Acanthamoeba* (host) and *Tetrahymena* (non-host, predator of particles). Thus, *A. castellanii* (900,000 cells) and *Tetrahymena* (100,000 cells) were added simultaneously to the same flask, then infected by Tupanvirus or mimivirus at an MOI of 10 and observed for 12 days. One flask per observation day was prepared. At days 4 and 8, we added 500 µL of fresh medium (50% PYG and 50% PAS) and 100,000 *A. castellanii*, the permissive host. Each day post-infection, the corresponding flask was collected and subjected to titration as previously described. The same experiment was carried out by pre-treating *Tetrahymena* (8 h before infection) with 20 µg/ml of geneticin as negative controls.

Analysis involving Tupanvirus 18 S rRNA intronic region. All analyses involving the genomic environment of copies 1 and 2 were conducted based on the annotation of Tupanvirus. In the best-hits evaluation, the core sequences of copies 1 and 2 were used for nucleotide BLAST analysis using blastn. The resulting 100 best hits were tabulated, and the information was used to construct diagrams. For the phylogenetic analysis, the sequences of these best hits were also aligned, using Clustal W in the Mega 7.0 software, and constructed using the maximum likelihood evolution method of 1,000 replicates. To analyze subjacent regions of the core sequence of 18 S rRNA intronic regions in the *Mimiviridae* family, one member of the lineages A (HQ336222.2), B (JX962719.1) and C (JX885207.1) was chosen and analyzed. The expression of both copies was checked using fluorescence in situ hybridization (FISH) and qPCR. For this, *A. castellanii* cells were infected with Tupanvirus with a multiplicity of infection of 5 and collected at 30 min and at 6 and 12 h post-infection. As a control, *A. castellanii* cells were also incubated with PAS alone and collected. At the indicated times, cells and the supernatant were collected and centrifuged at 800 × g for 10 min. For FISH, the pellet was resuspended in 200 µL of PAS, submitted to cytospin and the cells were fixed in cold methanol for 5 min. Specific probes targeting the 18 S rRNA of *A. castellanii* (5′-TTCACGGTAAACGATCTGGGCC-3′-fluorophore Alexa 488), copy 1 RNA (5′-AGTGGAACTCGGGTATGGTAAAA-3′-fluorophore Alexa 555) and copy 2 RNA (5′-GGCCAAGCTAATCACTTGGG-3′-fluorophore Alexa 555) were diluted and applied at 2 µM in hybridization buffer (900 mM NaCl, 20 mM Tris/HCL, 5 mM EDTA, 0.01% SDS, 10–25% deionized-formamide in distilled-H_2O). The hybridization buffer containing the probes was added to the slides and the hybridization was carried out at 46 °C overnight in a programmable temperature-controlled slide-processing system (ThermoBrite StatSpin, IL, USA). Post-hybridization washes consisted of 0.45–0.15 M NaCl, 20 mM Tris/HCL, 5 mM EDTA, and 0.01% SDS at 48 °C for 10 min. Slides were analyzed using a DMI6000B inverted research microscope (Leica, Wetzlar, Germany). To qPCR the pellet of infected cells was also washed with PAS and then used for total RNA extraction using the RNeasy mini kit (Qiagen, Venlo, Netherlands). The extracted RNA was treated with the Turbo DNA-free kit (Invitrogen, CA, USA) and then used as a template in reverse transcription (RT) reactions carried out using SuperScript Vilo (Invitrogen, CA, USA). The resultant cDNA was used as a template for quantitative real-time PCR assays using the QuantiTect SYBr Green PCR Kit (Qiagen RNA extraction Kit, Hilden, Germany) and targeting copies 1 (primers 5′-GCATCAA GTGCCAACCCATC-3′ and 5′-CTGAAATGGGCAATCCGCAG-3′) and 2 (primers 5′-CCAAGTGATTAGCTTGGCCATAA-3′ and 5′-CGGGAAGTCCCTA AAGCTCC-3′) of the intergenic18S rRNA region in TPV. To normalize the results, primers targeting the GAPDH housekeeping gene of *Acanthamoeba* (primers 5′-GTCTCCGTCGTCGATCTCAC-3′ and 5′-GCGGCCTTAATCTCGTCGTA-3′) were also used. qPCR assays were performed in a BioRad Real-Time PCR Detection System (BioRad) using the following thermal conditions for all genes: 15 min of pre-incubation at 95 °C followed by 40 amplification cycles of 30 s at 95 °C, 30 s at 60 °C and 30 s at 72 °C. The results were analyzed using the relative quantification methodology of $2^{(-\Delta\Delta ct)}$.

Investigation of the nature of ribosomal RNA shutdown. To investigate the shutting down of the host rRNA and verify whether this phenomenon was related to the canonical ribophagy/autophagy process, tests using two acidification and lysosome-vesicle fusion inhibitors (chloroquine and bafilomycin A) were performed. The pH of infected cells and the effect of Atg8-2 silencing on shutdown were also tested. For the inhibitor assays, 5×10^5 *A. castellanii* cells cultured in PYG medium were infected with Tupanvirus or mimivirus at a multiplicity of infection of 100 and incubated at 32 °C. At 1 h post-infection, chloroquine (Sigma-Aldrich, MO, USA) at a final concentration of 100 µM or bafilomycin A (Sigma-Aldrich, MO, USA) at a final concentration of 10 nM was added to the infected cell suspensions. As a control, *A. castellanii* cells not infected were also treated with these inhibitors under the same conditions. After 3 and 9 h post-infection, cells and the supernatant were collected and centrifuged at 800 × *g* for 10 min. The supernatant was discarded, and the pellet was submitted to RNA extraction (Qiagen RNA extraction Kit, Hilden, Germany). From the extracted RNA, 10 µL of each sample was electrophoresed in 1.5% agarose gel with TBE buffer and run at 135 V, and 14 µL was submitted to reverse transcription to measure the amoebal 18 S rRNA levels by qPCR as previously described. To investigate the acidification caused by Tupanvirus or mimivirus infection, *A. castellanii* cells were also submitted to the same pattern of infection and treatment with bafilomycin A, as previously described. In addition, 1 h before the collection time, the cells were incubated with LysoTracker Red DND-99 (Thermo Fisher Scientific, Massachusetts, United States) at a final concentration of 75 nM. After 9 h post-infection, cells and the supernatant were collected and centrifuged at 800 × *g* for 10 min. The supernatant was discarded, and the pellet was resuspended in 1 mL of PAS medium containing only bafilomycin A (10 nM). A total of 20 µL of this suspension was added to glass slides and cover slipped. Analyses were performed using a confocal microscope (Zeiss, Jena, Germany). For gene silencing, small interfering RNA (siRNA) targeting the Atg8-2 gene of *A. castellanii* was synthesized by Eurogentec (Liège, Belgium) based on the cDNA sequence of the gene. The siRNA duplex with sense (5′-GAACUC AUGUCGCACAUCUTT-3′) and anti-sense (5′-AGAUGUGCGACAUGAGU UCTT-3′) sequences was used. The siRNA tagged with a fluorescence dye was transfected onto *A. castellanii* trophozoites at a density of 1×10^6 cells. The control of transfection was performed using fluorescence microscopy. The biological effect of siRNA was check by qPCR and by the observation of the inhibition of acanthamoebal encystment, which is dependent on Atg8-2. Finally, modifications of *A. castellanii* nucleus/nucleolus structure after infection with Tupanvirus and mimivirus were investigated. A total of 10^6 cells were infected with Tupanvirus or mimivirus at an MOI of 10, stained by haemacolour and treated with SYTO RNASelect Green Fluorescent cell stain (Invitrogen, USA) following the manufacturer's instructions. After 9 h.p.i., cells were observed under an immunofluorescence microscope to observe modifications to the nucleus /nucleolus of infected and control cells. In parallel, this preparation was submitted to electron microscopy.

Data availability. The Tupanvirus genome sequences have been deposited in GenBank under accession codes KY523104 (soda lake) and MF405918 (deep ocean). Proteomic data have been deposited in PRIDE archive under accession code PXD007583. All other data supporting the findings of this study are available within the article and its Supplementary Information, or from the corresponding author upon reasonable request.

Received: 3 October 2017 Accepted: 23 January 2018
Published online: 27 February 2018

References

1. Raina, M. & Ibba, M. tRNAs as regulators of biological processes. *Front. Genet.* **5**, 171 (2014).
2. Fournier, G. P., Andam, C. P., Alm, E. J. & Gogarten, J. P. Molecular evolution of aminoacyl tRNA synthetase proteins in the early history of life. *Orig. Life Evol. Biospheres.* **41**, 621–632 (2011).
3. Korobeinikova, A. V., Garber, M. B. & Gongadze, G. M. Ribosomal proteins: structure, function, and evolution. *Biochemistry* **77**, 562–574 (2012).
4. La Scola, B. et al. A giant virus in amoebae. *Science* **299**, 2033 (2003).
5. Abrahao, J. et al. The analysis of translation-related gene set boosts debates around origin and evolution of mimiviruses. *PLoS Genet.* **13**, e1006532 (2017).
6. Schulz, F. et al. Giant viruses with an expanded complement of translation system components. *Science* **356**, 82–85 (2017).
7. Sorokin, D. Y. et al. Microbial diversity and biogeochemical cycling in soda lakes. *Extremophiles* **18**, 791–809 (2014).
8. Xiao, C. et al. Structural studies of the giant mimivirus. *PLoS Biol.* **7**, e92 (2009).
9. Ageno, M., Donelli, G. & Guglielmi, F. Structure and physico-chemical properties of bacteriophage G. II. The shape and symmetry of the capsid. *Micron* **4**, 376–403 (1973).
10. Philippe, N. et al. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* **341**, 281–286 (2013).
11. Legendre, M. et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. USA* **111**, 4274–4279 (2014).
12. Legendre, M. et al. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proc. Natl. Acad. Sci. USA* **112**, E5327–E5335 (2015).
13. Reteno, D. G. et al. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* **89**, 6585–6594 (2015).
14. Andreani, J. et al. Cedratvirus, a double-cork structured giant virus, is a distant relative of Pithoviruses. *Viruses* **8**, E300 (2016).
15. Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L. & Raoult, D. Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga Mimivirus. *PLoS ONE* **2**, e328 (2007).
16. Chaikeeratisak, C. et al. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* **355**, 194–197 (2017).
17. Forterre, P. The virocell concept and environmental microbiology. *ISME J.* **7**, 233–236 (2013).
18. Antwerpen, M. H. et al. Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. *Genome Announc.* **3**, e00136–15 (2015).

19. Raoult, D. The post-Darwinist rhizome of life. *Lancet* **375**, 104–105 (2010).
20. Suhre, K., Audic, S. & Claverie, J. M. Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proc. Natl. Acad. Sci. USA* **102**, 14689–14693 (2005).
21. Fischer, M. G., Allen, M. J., Wilson, W. H. & Suttle, C. A. Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc. Natl Acad. Sci. USA* **107**, 19508–19513 (2010).
22. Raoult, D. et al. The 1.2-megabase genome sequence of Mimivirus. *Science* **19**, 1344–1350 (2004).
23. Arslan, D., Legendre, M., Seltzer, V., Abergel, C. & Claverie, J. M. Distant mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl Acad. Sci. USA* **108**, 17486–17491 (2011).
24. Assis, F. L. et al. Pan-genome analysis of Brazilian lineage A amoebal mimiviruses. *Viruses* **7**, 3483–3499 (2015).
25. Gonzales-Flores, J. N., Shetty, S. P., Dubey, A. & Copeland, P. R. The molecular biology of selenocysteine. *Biomol. Concepts* **4**, 349–365 (2013).
26. Csaba, G. Lectins and Tetrahymena—A review. *Acta Microbiol. Immunol. Hung.* **63**, 279–291 (2016).
27. Jah, S. et al. Trans-kingdom mimicry underlies ribosome customization by a poxvirus kinase. *Nature* **546**, 651–655 (2017).
28. Claverie, J. M. Viruses take center stage in cellular evolution. *Genome Biol.* **7**, 110 (2006).
29. Filée, J. Genomic comparison of closely related Giant Viruses supports an accordion-like model of evolution. *Front. Microbiol.* **6**, 593 (2015).
30. Smith, J. E. The ecology and evolution of microsporidian parasites. *Parasitology* **136**, 1901–1914 (2009).
31. Nunes, A. & Gomes, J. P. Gomes Evolution, phylogeny, and molecular epidemiology of Chlamydia. *Infect. Genet. Evol.* **23**, 49–64 (2014).
32. Raoult, D. & Forterre, P. Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* **6**, 315–319 (2008).
33. Yutin, N. et al. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* **0**, 38–52 (2014).
34. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).
35. Rodrigues, R. A. et al. Mimivirus fibrils are important for viral attachment to the microbial world by a diverse glycoside interaction repertorie. *J. Virol.* **89**, 1182–1189 (2015).

Acknowledgements

We thank our colleagues from URMITE and Laboratório de Vírus of Universidade Federal de Minas Gerais for their assistance, particularly Julien Andreani, Jean-Pierre Baudoin, Gilles Audoly, Amina Cherif Louazani, Lina Barrassi, Priscilla Jardot, Eric Chabrières, Philippe Decloquement, Nicholas Armstrong, Said Azza, Emeline Baptiste, Claudio Bonjardim, Paulo Ferreira, Giliane Trindade and Betania Drumond. In addition, we thank the Méditerranée Infection Foundation, Centro de Microscopia da UFMG, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPEMIG (Fundação de Amparo à Pesquisa do estado de Minas Gerais) for their financial support. We thank Petrobras for the collection of sediments from ocean. This work was also supported by the French Government under the « Investissements d'avenir » (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03). J.A., B.R. and E.K. are CNPq researchers. B.L.S., J.A., L.S., P.C., and E.G.K. are members of a CAPES-COFECUB project.

Author contributions

D.R., B.L.S., J.S.A., A.L., P.C., E.G.K, and E.G. designed the study and experiments. L.S., J. S.A, J.B.K., R.R., L.S., L.S.S., T.A., P.C., F.A. P.B., M.A., I.B., B.R., A.L., and H.S. performed sample collection, virus isolation, experiments and/or analyses. D.R., B.L.S., A.L., J.S.A., P.C., G.K., R.R., L.S., and L.S.S. wrote the manuscript. All authors approved the final manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-03168-1.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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

jove Journal of Visualized Experiments www.jove.com

Video Article

Single Cell Micro-aspiration as an Alternative Strategy to Fluorescence-activated Cell Sorting for Giant Virus Mixture Separation

Dehia Sahmi-Bounsiar[1], Paulo Victor de Miranda Boratto[2], Graziele Pereira Oliveira[2], Jacques Yaacoub Bou Khalil[1], Bernard La Scola[1,2], Julien Andreani[2]

[1]Institut Hospitalo-Universitaire (IHU) - Méditerranée Infection

[2]Microbes, Evolution, Phylogeny and Infection (MEΦI), Aix-Marseille Université UM63, Institut de Recherche pour le Développement IRD 198, Assistance Publique - Hôpitaux de Marseille (AP-HM)

Correspondence to: Bernard La Scola at bernard.la-scola@univ-amu.fr, Julien Andreani at miaguiabidou@gmail.com

URL: https://www.jove.com/video/60148
DOI: doi:10.3791/60148

Keywords: Biology, Issue 152, single cell micro-aspiration, cloning, giant viruses, Faustovirus, co-culture, *Vermamoeba vermiform*, flow cytometry, FACS

Date Published: 10/27/2019

Citation: Sahmi-Bounsiar, D., Boratto, P.V., Oliveira, G.P., Bou Khalil, J.Y., La Scola, B., Andreani, J. Single Cell Micro-aspiration as an Alternative Strategy to Fluorescence-activated Cell Sorting for Giant Virus Mixture Separation. *J. Vis. Exp.* (152), e60148, doi:10.3791/60148 (2019).

Abstract

During the amoeba co-culture process, more than one virus may be isolated in a single well. We previously solved this issue by end point dilution and/or fluorescence activated cell sorting (FACS) applied to the viral population. However, when the viruses in the mixture have similar morphologic properties and one of the viruses multiplies slowly, the presence of two viruses is discovered at the stage of genome assembly and the viruses cannot be separated for further characterization. To solve this problem, we developed a single cell micro-aspiration procedure that allows for separation and cloning of highly similar viruses. In the present work, we present how this alternative strategy allowed us to separate the small viral subpopulations of Clandestinovirus ST1 and Usurpativirus LCD7, giant viruses that grow slowly and do not lead to amoebal lysis compared to the lytic and fast-growing Faustovirus. Purity control was assessed by specific gene amplification and viruses were produced for further characterization.

Video Link

The video component of this article can be found at https://www.jove.com/video/60148/

Introduction

Nucleocytoplasmic large DNA viruses (NCLDV) are extremely diverse, defined by four families that infect eukaryotes[1]. The first described viruses with genomes above 300 kbp were *Phydcodnaviridae*, including *Paramecium bursaria Chlorella virus 1* PBCV1[2]. The isolation and the first description of Mimivirus, showed that the size of viruses doubled in terms of both the size of the particle (450 nm) and the length of the genome (1.2 Mb)[3]. Since then, many giant viruses have been described, usually isolated using an amoeba co-culture procedure. Several giant viruses with different morphologies and genetic contents can be isolated from *Acanthamoeba* sp. cells, including Marseilleviruses, Pandoraviruses, Pithoviruses, Mollivirus, Cedratviruses, Pacmanvirus, Tupanvirus, and recently Medusavirus[4,5,6,7,8,9,10,11,12,13,14,15,16,17]. In parallel, the isolation of *Vermamoeba vermiformis* allowed the isolation and description of the giant viruses Faustovirus, Kaumoebavirus, and Orpheovirus[18,19,20]. Other giant viruses were isolated with their host protists, such as *Cafeteria roenbergensis*[21], *Aureococcus anophagefferens*[22], *Chrysochromulina ericina*[23], and *Bodo saltans*[24]. All of these isolations were the result of an increasing number of teams working on isolation and the introduction of high throughput strategy updates[25,26,27,28], such as the improvement of the co-culture system with the use of flow cytometry.

In 2016, we used a strategy associating co-culture and flow cytometry to isolate giant viruses[27]. This strategy was developed to increase the number of samples inoculated, to diversify protists used as cell supports, and to quickly detect the lysis of the cell support. The system was updated by adding a supplemental step to avoid preliminary molecular biology identification and quick detection of an unknown viral population as in the case of Pacmanvirus[29]. Coupling flow cytometry to cell sorting allowed for separation of a mixture of Mimivirus and Cedratvirus A11[30]. However, we later encountered the limitations of the separation and detection of these viral subpopulations by flow cytometry. After sequencing, when we assembled the genomes of Faustovirus ST1[25] and Faustovirus LCD7 (unpublished data), we surprisingly found in each assembly two supplemental genomes of two novel viruses not identified in public genome databases. However, neither flow cytometry nor transmission electronic microscopy (TEM) showed that the amoebaes were infected by two different viruses, Clandestinovirus ST1 and Usurpativirus LCD7. We designed specific PCR systems to amplify Faustovirus, Usurpativirus, and Clandestinovirus markers respectively based on their genomes; our purpose was to have PCR-based systems that enable verification of the purity of the viruses being separated. However, end-point dilution and flow cytometry failed to separate them. The isolation of this single viral population was difficult because neither the morphology nor replicative elements of Clandestinovirus and Usurpativirus populations have been characterized. We detected only one viral population by flow cytometry due to the overlapping of the two populations (tested after the effective separation). We tried to separate them using single particle sorting on 96-well plates, but we did not observe any cytopathic effects, and we detected neither Clandestinovirus nor Usurpativirus by PCR

amplification. Finally, it was only the combination of end point dilution followed by single amoeba micro-aspiration that enabled separation of these two low-abundance giant viruses from Faustoviruses. This method of separation is the object of this article.

Protocol

1. Amoeba Culture

1. Use *Vermamoeba vermiformis* (strain CDC19) as a cell support.
2. Add 30 mL of protease-peptone-yeast extract-glucose medium (PYG) (**Table 1**) and 3 mL of the amoebae at a concentration of 1 x 10^6 cells/mL in a 75 cm^2 cell culture flask.
3. Maintain the culture at 28 °C.
4. After 48 h, quantify the amoebae using counting slides.
5. To rinse, harvest the cells at a concentration of 1 x 10^6 cells/mL and pellet the amoebae by centrifugation at 720 x *g* for 10 min. Remove the supernatant and resuspend the pellet in the appropriate volume of starvation medium to obtain 1 x 10^6 cells/mL (**Table 1**).

2. Propagation of Stock Virus in Amoebae

NOTE: Before dilution, it is important to culture the stock sample to obtain enough fresh culture, then proceed to filtration.

1. Use 1 x 10^6 of amoeba culture in starvation medium.
2. Inoculate the mixture of viruses issued from the stock solution (Faustovirus/Usurpativirus LCD7 or Faustovirus/Clandestinovirus ST1) after the co-culture process on the cell support at a multiplicity of infection (MOI) of 0.01.
   NOTE: The MOI is important to reduce the abundance of the major viral population and the number of infected cells.
3. Incubate at 30 °C until cytopathic effects (CPE) are induced, such as amoebal rounding or lysis, approximately 10 to 14 h after infection.
4. Collect the media and filtrate through a 5 µm filter to remove cellular debris.

3. End-point dilution

1. Perform a serial dilution (10^{-1} to 10^{-11}) of the viral sample in starvation medium (**Table 1**).
2. Inoculate 2 mL of 1 x 10^6 *Vermamoeba vermiformis* contained in each Petri dish with 100 µL of the mixture inoculum.
3. Place the Petri dishes into a sealable plastic bag at 30 °C.
4. Begin observing the Petri dishes with inverted optical microscopy at 6 h postinfection and check cell morphology every 4 to 8 h.
5. At the appearance of the cytopathic effect characterized by rounding cells, begin the single cell micro-aspiration process.

4. Single Cell Micro-aspiration

1. **Prepare the host.**
   NOTE: This preparation is made for the release of infected single cells to a fresh cell support.
   1. Treat the amoebae in the culture with an antimicrobial agent containing 10 µg/mL of vancomycin, 10 µg/mL of imipenem, 20 µg/mL of ciprofloxacin, 20 µg/mL of doxycycline, and 20 µg/mL of voriconazole. This mixture is used to avoid bacterial and fungal contamination.
      NOTE: The procedure takes place on a bench outside the microbiological safety station. Add 2 mL of amoebae concentrated at 1 x 10^6 cells/mL each into 15 Petri dishes. For amoeba adherence, incubate the culture at 30 °C for 30 min.

2. Select the Petri dish used for the micro-aspiration from the limit dilution according to the following criteria: 1) absence of any visible contamination by fungal and bacterial agents, 2) evidence of cytopathic effect of amoebae due to the viruses, and 3) prelysis and rounding phase of the amoebae (to avoid aspiration of viral particles).
3. **Set up a workstation with the following materials (see Figure 1A,B):**
   Micromanipulator, which allows microcapillary positioning;
   Manual control pressure device, used to aspirate and release the cells into the microcapillary;
   Inverted microscope;
   Plug and play motor modules;
   Camera;
   Computer module to visualize manipulation and take pictures.
4. Choose a microcapillary (see **Figure 1C**).
   NOTE: The size of the cells, the deformation and adhesion of their membranes to the surfaces, and the cellular motility can impact the smooth progress of the micro-aspiration. The microcapillary diameter can be precisely chosen and adapted to specific cell types depending on their sizes and methods of aspiration. A microcapillary of 20 µm inner diameter was used to aspirate a rounding amoeba (diameter ~10 µm). This allows the upkeep of an internal position and an easy release of the cell.
5. **Mount the system.**
   1. Fix the operating angle of the gripping system on the motorized module at 45°.
   2. Perform a double installation, first on the gripping system, and then on the microcapillary.
   3. Focus on the cells after running a few drops of oil through the microcapillary.
      NOTE: The mineral oil with biological compatibility is supplied by the device.
   4. Complete mounting following manufacturer's recommendations.

6. **Clone cells (see Figure 2A,B).**
   NOTE: This procedure is similar to the one described by Fröhlich and König[31].

1. Place the Petri dish containing 2 mL of infected amoebae under the microscope.
2. Focus first on the cells, and then on the microcapillary immersed in the culture.
3. Pick a rounded single cell and bring the microcapillary closer to the micromanipulator.
4. Exert soft aspiration with manual pressure control on the cell, taking it inside the microcapillary. Remove the single cell from the first sample and release in the cellular support, then incubate it at 30 °C.
5. Conduct daily observations with an inverted optical microscope to observe the appearance of the cells and to monitor the emergence of the cytopathic effect.

5. PCR Screening

NOTE: Following step 4, a systematic screening by PCR is crucial to confirm the separation. In both Usurpativirus/Faustovirus and Clandestinovirus/Faustovirus, the design and application of the specific primer and probe systems were done using Primer-BLAST online[32] (**Table 2**).

1. Extract DNA from a part of the positive culture samples (i.e., where a cytopathic effect is observed), using an automated extraction system according to the manufacturer's protocol.
2. Use appropriately designed primers.
   NOTE: Here we designed primers to amplify core genes annotated as RpB2 (Faustovirus), LCD7 major capsid protein (Usurpatvirus) and minor capsid protein (Clandestinovirus)
3. **Perform standard PCR using a thermocycler.**
   1. Carry out 20 µL PCR reactions with 50 µM of each primer (**Table 2**), 1x Master Mix, and RNase free water.
   2. Activate the Taq DNA polymerase for 5 min at 95 °C, then follow with 45 cycles of 10 s denaturation at 95 °C, annealing of the primers for 30 s at 58 °C, and extension for 30 s at 72 °C.
4. Run the PCR products on a 1.5% agarose gel, stain with DNA gel stain (**Table of Materials**), and visualize with UV.

6. Virus Production and Purification

1. Put the rest of the Petri dish culture back in a small flask.
2. For the virus production, prepare 15 flasks of 145 cm^2, containing 40 mL of *Vermamoeba vermiformis* in starvation medium and 5 mL of the isolated virus already transferred from the Petri dish to small flasks.
3. Treat with the same antibiotic and antifungal mixture used in step 4.1.
4. Incubate at 30 °C. Observe every day with inverted optical microscopy.
5. After the complete infection, pool all flasks. Use a 0.45 µm filter to eliminate debris.
6. Ultracentrifuge all supernatants at 50,000 x *g* for 45 min.
7. After centrifugation, remove the supernatant from each tube by aspiration and resuspend the pellet in 1 mL of phosphate buffered saline (PBS).
8. Purify the virus produced using 25% sucrose (27.5 g sucrose in 100 mL of PBS, sterilized by filtration).
9. Centrifuge 8 mL of sucrose and 2 mL of the viral suspension at 80,000 x *g* for 30 min. Resuspend the viral pellet in 1 mL of PBS. Store it at -80 °C.

7. Negative Staining and Transmission Electron Microscopy

NOTE: Bou Khalil et al. previously published this protocol[27].

1. Deposit 5 µL of the lysis supernatant onto the glow-discharged grid. Leave for approximately 20 min at room temperature.
   NOTE: The glow-discharge allows us to obtain a hydrophilic grid by plasma application.
2. Dry the grid carefully and deposit a small drop of 1% ammonium molybdate on it for 10 s. Leave the grid to dry for 5 min.
3. Proceed to electron microscopy observations at 200 keV.

8. Characterization of Clandestinovirus ST1 and Usurpativirus LCD7

1. Characterize pure populations of Clandestinovirus ST1 and Usurpativirus LCD7 using genome sequencing, genome assembly, bioinformatics analyses, and study of their replicative cycle as we have done for other viruses[10,20,29].

Representative Results

Single cell micro-aspiration is a micromanipulation process optimized in this manuscript (**Figure 1**). This technique enables capture of a rounded, infected amoeba (**Figure 2A**) and its release in a novel plate containing uninfected amoebae (**Figure 2**). It is a functional prototype that applies to the co-culture system and has successfully isolated non-lytic giant viruses. This approach was used for the first time in the field of giant viruses and made it possible to isolate two new low-abundance giant viruses. We named the new viruses Clandestinovirus ST1 and Usurpativirus LCD7, that were in low abundance compared to the high abundance Faustoviruses. In order to analyze the viral presence in each plate after the micro-aspiration procedure, we applied PCR on the 15 micro-aspirations. We observed pure Usurpativirus (presence of Usurpativirus LCD7 [+] and absence of Faustovirus [–]) only in clone 7 (**Figure 3**). The purity of the clone was confirmed by PCR with specific protocols targeting Clandestinovirus ST1 and Usurpativirus LCD7 (**Table 2**). Electron microscopy revealed the appearance of Clandestinovirus ST1 and Usurpativirus LCD7 (**Figure 4**), which have a typical icosahedral morphology without fibrils and an icosahedral capsid of about 250

nm, respectively. After the confirmation of the production purity, the clonal virus was produced and purified for whole genome sequencing and further characterization, especially transmission electronic microscopy (TEM) for multiplication cycle studies. We confirmed the overlapping of populations (Faustovirus/novel virus) by flow cytometry (**Figure 5**).

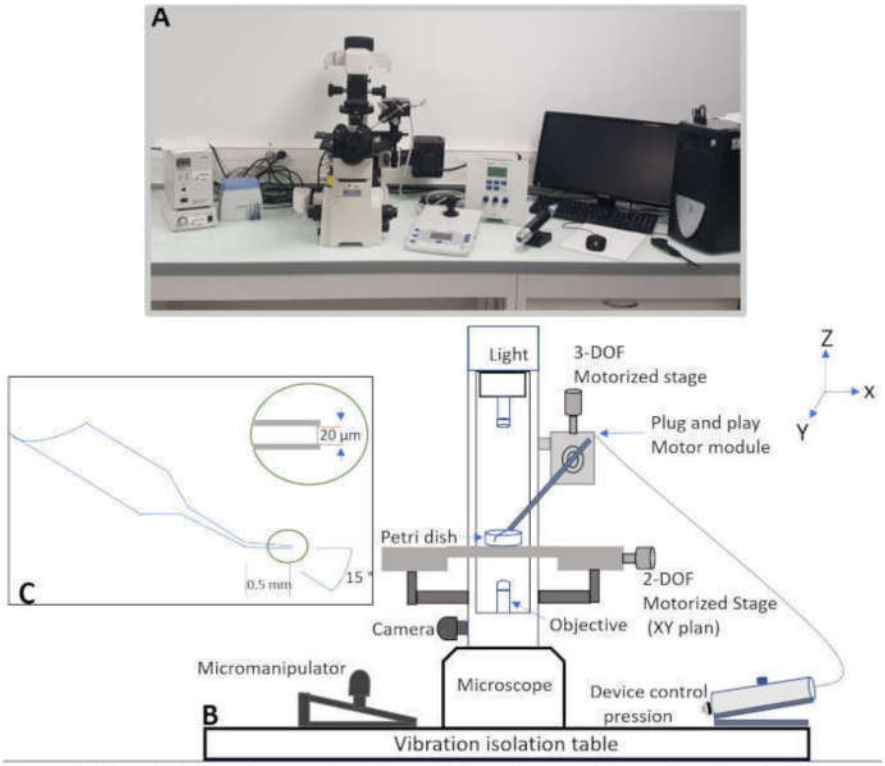


Figure 1: Materials for micromanipulation. (**A**) Actual setup of the workstation. (**B**) Schematic illustration of the workstation's components. (**C**) Schematic illustration of the microcapillary. Please click here to view a larger version of this figure.

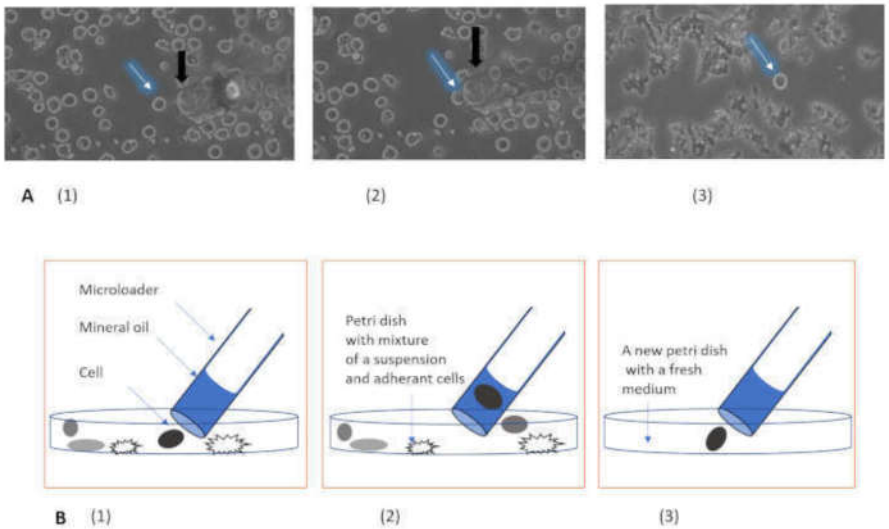


Figure 2: Micro-aspiration steps. (**A**) Single cell isolation procedure (zoom x40). (**B**) Schematic illustration of the different steps of single cell aspiration: 1) Localization of the cell. 2) Aspiration of the cell. 3) Release step. The black arrows show the microcapillary and the white ones show the single cell. Please click here to view a larger version of this figure.

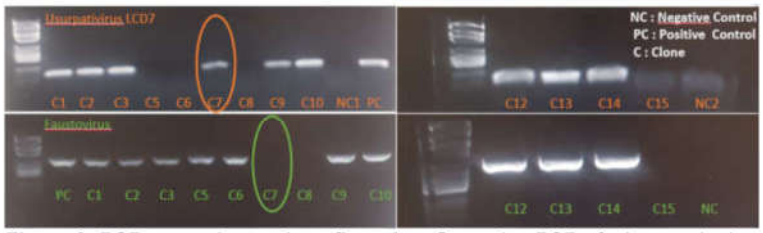


Figure 3: PCR screening and confirmation. Screening PCR of micro-aspirations carried out for Usurpativirus LCD7 and Faustovirus. Presence of Usurpativirus LCD7 DNA (+) and absence of Faustovirus DNA (–) observed for clone 7 after micro-aspiration procedure. Please click here to view a larger version of this figure.

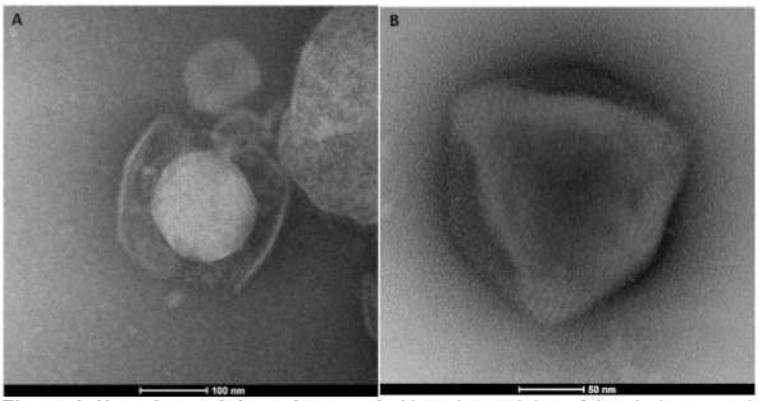


Figure 4: Negative staining micrograph. Negative staining of the viral suspension, showing pure Clandestinovirus ST1 (**A**) and Usurpativirus LCD7 (**B**). Please click here to view a larger version of this figure.

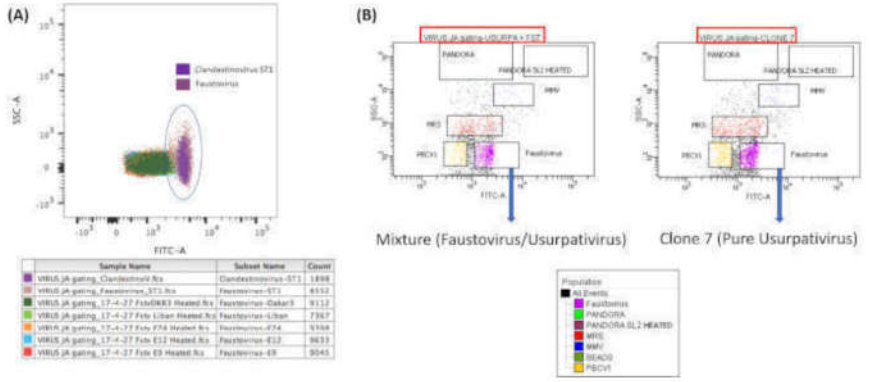


Figure 5: Representative gate plots. (**A**) Superimposition of each single viral population previously stained with fluorescent molecular probes for virus DNA labelling (following previously described protocol[26,30]). The left part of the picture shows the superposition of five strains of Faustovirus (dark green, light green, orange, red, and blue). On the right, the superposition of Faustovirus ST1 and Clandestinovirus ST1 (light purple and dark purple) is visible. (**B**) The presence of two viral population of Faustovirus and Usurpativirus in the same gate as Faustovirus (left). A dot plot of pure Usurpativirus shows the same gate defined previously for Faustovirus (right). Please click here to view a larger version of this figure.

PYG composition	quantities
Proteose peptone	20 g
Yeast extract	20 g
$MgSO_4 \cdot 7H_2O$	0.980 g
$CaCl_2$	0.059 g
Citrate sodium. Dihydrate	1 g
Fe (NH_4) 2(SO_4) 2 x $6H_2O$	0.02 g
Glucose	18 g
Distilled water	1 L
Adjust pH at 6.8 with HCl or KOH	
Autoclave 15 min at 121 °C	
Starvation medium	**quantites**
Yeast extract	2 g
Glucose	18 g
Fe (NH_4) 2(SO_4) 2 x $6H_2O$	0.02 g
PAS (detailed below)	1 L
PAS solution A	**quantites**
KH_2PO_4	0.136 g
Na_2HPO_4	0.142 g
PAS solution B	**quantites**
$MgSO_4 \cdot 7H_2O$	4.0 mg
$CaCl_2 \cdot 2H_2O$	4.0 mg
NaCl	0.120 g
10 mL each of solution A and B are added into 1 L of distilled water.	

Table 1: Composition of Culture Media.

Virus target	Target gene	Foward primer (5->3)	Reverse primer (5->3)
Faustovirus RpB2	RpB2	CWCAATCHGCYGATACHGGTGA	TGATTWGCYAATGGNGCYGC
Usurpatvirus LCD7 Major capsid gene	Major capsid gene	GGGCAAGAAGCTCCAAGCTA	GGGTTGAGGAGGAGTCAACG
Clandestinovirus ST1 Minor capsid gene	Minor capsid gene	AAAATGAACGCGTTGGAGGC	ACCGGCGAATGTTCCTATGG

Table 2: Primmer Sequences Used for the PCR.

Discussion

The duration of the single cell micro-aspiration handling and its good functioning is operator-dependent. The different steps of the experiment require precision. The use of the micromanipulation components of the workstation must be under constant control by observing the process of micro-aspiration and the release of the cell. The follow-up by microscopic observation is necessary for capture and transfer of a cell. An experienced operator can take 1 to 2 h to isolate 10 cells and retransfer them one by one depending on the abundance of viruses to be isolated. The number of manipulations can vary. We advise beginning with 10 manipulations. By definition, we do not know the intrinsic characteristics of the unknown virus. Thus, the use and development of different strategies for isolation are necessary to optimize the success of separating these viruses.

The micro-aspiration process is performed under unsterile conditions (on a bench outside the microbiological safety station) and thus restricts its usage and prevents its application for the study of human pathogens. Therefore, the use of a mixture of antibiotics and antifungals to limit contaminants is mandatory. Another limitation of the method is that it can only be performed on microorganisms observable by light microscopy on which the micromanipulation components were mounted, thus, conceptually eliminating any work on microorganisms not observable by light microscopy. However, we were able to separate giant viruses of about 200 nm which remain invisible under the light microscope by using an indirect strategy consisting of separating and cloning the infected hosts.

The development of single cell micro-aspiration for amoebal capture is a part of the development of population-sorting methods. Single cell micro-aspiration allowed us to isolate two new giant viruses, Clandestinovirus ST1 and Usurpativirus LCD7, with characteristics distinctive from

Faustoviruses. The highly similar morphological properties of both Clandestinovirus ST1 and Usurpativirus LCD7 to Faustovirus LCD7 and their difference of replication shows the limit of FACS, which is usually used in giant viruses sorting. It is represented by the superposition of two viral populations, Clandestinovirus ST1 and Faustovirus ST1 (**Figure 5A**) and also by the detection for Usupartivirus LCD7 and Faustovirus LCD7. However, with this new method, the entire manipulation is under visual control with careful monitoring of the cell and its integrity even after its release. The use of flow cytometry to directly sort infected amoebae could be an alternative solution to indirectly sort novel viruses. This method is usually followed by the screening of the plate in order to detect cytopathic effects or lysis. The presence of non-lytic viruses in the mix could represent a limit to amoebal sorting. However, this was not explored in the creation of this protocol and for these two novel isolations.

Beyond the application of the micromanipulation in the positioning and holding of cells in general and oocytes for intracytoplasmic sperm injection (ICSI)[33], single cell micro-aspiration has proven to be a practical method for the isolation of single prokaryotic cells observable under an optical microscope[31]. Other applications could be tested, including cell sorting on a morphological basis to have different pure samples from a mixed sample of microorganisms. The observable separation of microorganisms can be envisioned using the strategy described above.

Disclosures

All authors have nothing to disclose.

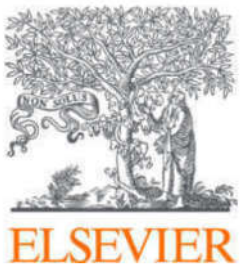
Acknowledgments

The authors would like to thank both Jean-Pierre Baudoin and Olivier Mbarek for their advice and Claire Andréani for her help in English corrections and modifications. This work was supported by a grant from the French State managed by the National Research Agency under the "Investissements d'avenir (Investments for the Future)" program with the reference ANR-10-IAHU-03 (Méditerranée Infection) and by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

References

1. Iyer, L. M., Aravind, L., Koonin, E. V. Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. *Journal of Virology.* **75** (23), 11720–11734 (2001).
2. Li, Y. et al. Analysis of 74 kb of DNA located at the right end of the 330-kb chlorella virus PBCV-1 genome. *Virology.* **237** (2), 360–377 (1997).
3. Scola, B. L. A Giant Virus in Amoebae. *Science.* **299** (5615), 2033 (2003).
4. Philippe, N. et al. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science.* **341** (6143), 281–286 (2013).
5. Antwerpen, M. H. et al. Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. *Genome Announcements.* **3** (2), e00136–15 (2015).
6. Legendre, M. et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proceedings of the National Academy of Sciences of the United States of America.* **111** (11), 4274–4279 (2014).
7. Legendre, M. et al. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proceedings of the National Academy of Sciences of the United States of America.* **112** (38), E5327–335 (2015).
8. Legendre, M. et al. Diversity and evolution of the emerging Pandoraviridae family. *Nature Communications.* **9** (1), 2285 (2018).
9. Aherfi, S. et al. A Large Open Pangenome and a Small Core Genome for Giant Pandoraviruses. *Frontiers in Microbiology.* **9**, 166 (2018).
10. Andreani, J. et al. Cedratvirus, a Double-Cork Structured Giant Virus, is a Distant Relative of Pithoviruses. *Viruses.* **8** (11), 300 (2016).
11. Rodrigues, R. A. L. et al. Morphologic and genomic analyses of new isolates reveal a second lineage of cedratviruses. *Journal of Virology.* JVI.00372–18 (2018).
12. Bertelli, C. et al. Cedratvirus lausannensis- digging into Pithoviridaediversity. *Environmental Microbiology.* (2017).
13. Levasseur, A. et al. Comparison of a modern and fossil Pithovirus reveals its genetic conservation and evolution. *Genome Biology and Evolution.* evw153 (2016).
14. Dornas, F. P. et al. A Brazilian Marseillevirus Is the Founding Member of a Lineage in Family Marseilleviridae. *Viruses.* **8** (3), 76 (2016).
15. Yoshikawa, G., Blanc-Mathieu, R. et al. Medusavirus, a novel large DNA virus discovered from hot spring water. *Journal of Virology.* JVI.02130–18 (2019).
16. Legendre, M. et al. Pandoravirus Celtis Illustrates the Microevolution Processes at Work in the Giant Pandoraviridae Genomes. *Frontiers in Microbiology.* **10**, 430 (2019).
17. Abrahão, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nature Communications.* **9** (1), 749 (2018).
18. Reteno, D. G. et al. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *Journal of Virology.* **89** (13), 6585–6594 (2015).
19. Bajrai, L. et al. Kaumoebavirus, a New Virus That Clusters with Faustoviruses and Asfarviridae. *Viruses.* **8** (12), 278 (2016).
20. Andreani, J. et al. Orpheovirus IHUMI-LCC2: A New Virus among the Giant Viruses. *Frontiers in Microbiology.* **8**, 779 (2018).
21. Fischer, M. G., Allen, M. J., Wilson, W. H., Suttle, C. A. Giant virus with a remarkable complement of genes infects marine zooplankton. *Proceedings of the National Academy of Sciences of the United States of America.* **107** (45), 19508–19513 (2010).
22. Moniruzzaman, M. et al. Genome of brown tide virus (AaV), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host-virus coevolution. *Virology.* **466-467**, 60–70 (2014).
23. Gallot-Lavallée, L., Blanc, G., Claverie, J.-M. Comparative genomics of Chrysochromulina Ericina Virus (CeV) and other microalgae-infecting large DNA viruses highlight their intricate evolutionary relationship with the established Mimiviridae family. *Journal of Virology.* JVI.00230–17 (2017).
24. Deeg, C. M., Chow, C.-E. T., Suttle, C. A. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. *eLife.* **7**, 235 (2018).

25. Boughalmi, M. et al. High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. *Environmental Microbiology.* **15** (7), 2000–2007 (2013).
26. Khalil, J. Y. B. et al. High-Throughput Isolation of Giant Viruses in Liquid Medium Using Automated Flow Cytometry and Fluorescence Staining. *Frontiers in Microbiology.* **7** (120), 26 (2016).
27. Bou Khalil, J. Y., Andreani, J., Raoult, D., La Scola, B. A Rapid Strategy for the Isolation of New Faustoviruses from Environmental Samples Using Vermamoeba vermiformis. *Journal of Visualized Experiments : JoVE.* (112), e54104–e54104 (2016).
28. Khalil, J. Y. B., Andreani, J., La Scola, B. Updating strategies for isolating and discovering giant viruses. *Current Opinion in Microbiology.* **31**, 80–87 (2016).
29. Andreani, J. et al. Pacmanvirus, a new giant icosahedral virus at the crossroads between Asfarviridae and Faustoviruses. *Journal of Virology.* JVI.00212–17 (2017).
30. Khalil, J. Y. B. et al. Flow Cytometry Sorting to Separate Viable Giant Viruses from Amoeba Co-culture Supernatants. *Frontiers in Cellular and Infection Microbiology.* **6**, e17722 (2017).
31. Fröhlich, J., König, H. New techniques for isolation of single prokaryotic cells. *FEMS Microbiology Reviews.* **24** (5), 567–572 (2000).
32. Ye, J. et al. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* **13** (1), 134 (2012).
33. Kimura, Y., Yanagimachi, R. Intracytoplasmic sperm injection in the mouse. *Biology of Reproduction.* **52** (4), 709–720 (1995).



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Science of the Total Environment 766 (2021) 142645

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Detection of SARS-CoV-2 RNA on public surfaces in a densely populated urban area of Brazil: A potential tool for monitoring the circulation of infected patients

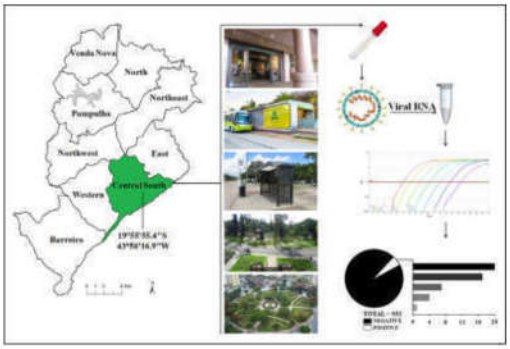
Jônatas Santos Abrahão [a,*], Lívia Sacchetto [a], Izabela Maurício Rezende [a], Rodrigo Araújo Lima Rodrigues [a,b,**], Ana Paula Correia Crispim [a], César Moura [a], Diogo Correa Mendonça [a], Erik Reis [a], Fernanda Souza [a], Gabriela Fernanda Garcia Oliveira [a], Iago Domingos [a], Paulo Victor de Miranda Boratto [a], Pedro Henrique Bastos Silva [a], Victoria Fulgêncio Queiroz [a], Talita Bastos Machado [a], Luis Adan Flores Andrade [a], Karine Lima Lourenço [a], Thaís Silva [c], Graziele Pereira Oliveira [a], Viviane de Souza Alves [a], Pedro Augusto Alves [c], Erna Geessien Kroon [a], Giliane de Souza Trindade [a], Betânia Paiva Drumond [a]

[a] Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[b] Federal University of Ouro Preto, Ouro Preto, Minas Gerais, Brazil
[c] René Rachou Institute – FIOCRUZ, Belo Horizonte, Minas Gerais, Brazil

HIGHLIGHTS

- COVID-19 has caused over 4.4 million cases in Brazil by the end of September 2020.
- The presence of SARS-CoV-2 in public surfaces can result in new infections.
- RT-qPCR was used for assessing the presence of viral genome in the environment.
- SARS-CoV-2 was detected especially in surfaces near hospitals and public squares.
- Constant surveillance of the virus in the environment is essential for disease control.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 July 2020
Received in revised form 22 September 2020
Accepted 23 September 2020
Available online 2 October 2020

Editor: Jay Gan

Keywords:
COVID-19
Coronavirus
Genome detection
Environment surveillance

ABSTRACT

The world is experiencing the worst global health crisis in recent decades since December/2019 due to a new pandemic coronavirus. The COVID-19 disease, caused by SARS-CoV-2, has resulted in more than 30 million cases and 950 thousand deaths worldwide as of September 21, 2020. Determining the extent of the virus on public surfaces is critical for understanding the potential risk of infection in these areas. In this study, we investigated the presence of SARS-CoV-2 RNA on public surfaces in a densely populated urban area in Brazil. Forty-nine of 933 samples tested positive (5.25%) for SARS-CoV-2 RNA, including samples collected from distinct material surfaces, including metal and concrete, and distinct places, mainly around hospital care units and public squares. Our data indicated the contamination of public surfaces by SARS-CoV-2, suggesting the circulation of infected patients and the risk of infection for the population. Constant monitoring of the virus in urban areas is required as a strategy to fight the pandemic and prevent further infections.

* Corresponding author.
** Correspondence to: R.A.L. Rodrigues, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.
E-mail addresses: jonatas.abrahao@gmail.com, jsa@icb.ufmg.br (J.S. Abrahão), rodriguesral07@gmail.com (R.A.L. Rodrigues).

https://doi.org/10.1016/j.scitotenv.2020.142645
0048-9697/

1. Introduction

Since late December 2019, the world is experiencing the worst global health crisis in recent decades due to the ongoing transmission of a novel coronavirus. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causing coronavirus disease 2019 (COVID-19) spread globally and affected several sectors, including those related to medicine, economics, and politics, among others (Cutler, 2020; Gorbalenya et al., 2020; Lai et al., 2020; Zhu et al., 2020). COVID-19 has already affected over 200 countries, resulting in more than 30.6 million cases and 950 thousand deaths worldwide as of September 21, 2020 (WHO, 2020). In Brazil, the first official case was registered on February 26, 2020, and more than 4,490,000 confirmed cases and 135,000 deaths had been registered across the country as of September 21, 2020 (WHO, 2020).

Belo Horizonte is the capital of Minas Gerais State, one of Brazil's most populous metropolitan regions (6 million inhabitants). The city recorded 39,379 confirmed cases and 1168 deaths due to COVID-19 as of September 21 (Secretaria Municipal de Saúde, 2020). The areas within Belo Horizonte where most of the deaths and confirmed cases have occurred correspond precisely to areas with public squares, bus stations/terminals, and hospital areas, i.e., where a large flow and concentration of people is commonly observed (Secretaria Municipal de Saúde, 2020).

Recent studies have identified the presence of SARS-CoV-2 RNA on different surfaces and environments inside hospitals, revealing the dynamics of viral dissemination within these places (Guo et al., 2020; Liu et al., 2020; Wang et al., 2020). SARS-CoV-2 remains viable on different types of surfaces, such as metal and plastic, for up to 72 h, depending on the type of surface material, and can remain infectious in aerosols for at least 3 h (Van Doremalen et al., 2020). Assessing the presence of the virus in the environment, objects, and surfaces in public areas is fundamental for understanding the risk of infection in the population. In addition, any information gained can be used by health managers to control population movements in these areas, as well as implement environmental disinfection measures.

2. Material and methods

2.1. Sample collection

We investigated the presence of SARS-CoV-2 RNA in a downtown area of Belo Horizonte between April and June 2020, in a total of nine days. Belo Horizonte's climate is classified as tropical with a dry season, with moderately hot and humid summers and dry and pleasant winters. The temperature is mild throughout the year, with averages ranging from 19 °C to 24 °C, with the annual compensated average of 22 °C and 1430 mm is the average annual rainfall (INMET, 2020). All samples were collected between 14:00 pm and 17:00 pm. Temperature was between 20 °C to 25 °C (average temperature 22 °C) and relative humidity varying between 36% to 83% (average humidity 54%). Environmental data was retrieved from Time and Date AS website (https://www.timeanddate.com/weather/brazil/belo-horizonte/). This region of the city (central-southeast: 19°55′55.4″S 43°56′16.9″W) has the highest concentration of hospitals and health units and one of the highest number of notified COVID-19 cases. Importantly, this part of the city also has a large number of people accessing public transportation and transportation facilities daily. A total of 933 samples were collected from eight different categories of places (Supplementary Table 1), including: a) 38 health care units (hospitals, medical centers, and emergency care units); b) 17 public squares; c) two public parks; d) one public market; e) six bus terminals; f) one shopping mall; g) 10 education centers (universities and schools); 8) 21 other public places, including banks, government departments, among others (Supplementary Table 1). Samples were collected from different sources including entrance doors, handrails, benches and tables, bus stops, and ground, and from distinct materials, including concrete, metal, rock, brickwork, and others. For the collection of environmental samples, swabs with sterile phosphate-buffered saline were vigorously rubbed on surfaces (10 cm^2) of the aforementioned local and objects. The swabs were then transferred to tubes containing transport solution (1 mL of guanidine isothiocyanate buffer, 4 M) and taken directly to the laboratory for testing.

2.2. RNA extraction and RT-qPCR

For each sample, 70 μL of transport solution containing a sample was submitted to nucleic acid extraction using the QIAmp Viral RNA Mini Kit (QIAGEN, Maryland, USA). Total RNA (5 μL) was used as a template for one-step qPCR (Promega, Wisconsin, USA) (in a final volume of 20 μL per reaction, GoTaq1-sept qPCR system, Promega), using primers and probes specific for the N1 and N2 regions of the SARS-CoV-2 genome (CDC, USA 2020). RNA extraction was performed in batches of 13 samples plus one negative control. Samples were considered positive when they presented amplification for N1 an N2 targets region, considering the threshold for cycle quantification value (Cq) of 40 (CDC, USA 2020). Since between the range of 37–40 Cq indicate minimal quantities of DNA, Cq ≥ 40 were considered negative. The results of RT-qPCR runs were manually inspected for the correction of baseline and threshold parameters whenever necessary due to heterogeneity in the amount of input RNA among different samples (Bustin et al., 2009). Negative (extraction control and non-template control) and positive controls (RNA extracted from inactivated SARS-CoV-2, kindly provided by Dr. Danielle Durigon and Dr. Edison Durigon, USP, Brazil and Dr. Rafael Elias Marques (Centro Nacional de Biociencias LNBio-CNPEM) were used. To confirm the results, all positive samples were submitted to a second round of RNA extraction and RT-qPCR. Quantification of viral RNA in the environmental samples was based on a standard curve generated from serial dilutions (1:10) of SARS-CoV-2 RNA and converted to genomic units per ten square centimeters of surface (the area we made the sampling for this study). To date, for PCR quantification (based on standard curve) we considered 1 SARS-CoV-2 plaque forming unit as 1 SARS-CoV-2 genomic unit. Quantification was based only on N1 target gene given the high efficiency of our standard curve (Supplementary Fig. 1). N2 target gene quantification was not performed due to a low efficiency achieved for the standard curve for this target. Nevertheless, the Cq values for N2 target for all positive samples are included in supplementary Table 2. SARS-CoV-2 RNA control was previously quantified as described elsewhere (de Almeida et al., 2020).

3. Results

A total of 49 samples (5.25%) were positive for the presence of SARS-CoV-2 RNA (Table 1 and Fig. 1). SARS-CoV-2 RNA was detected in 20 samples collected around health care units, corresponding to 40.8% of the total positive samples, with Ct values ranging from 23.3 to 37.7 (N1) and 22.2 to 39.4 (N2). These samples are distributed in 12 different health care units and were detected mainly at bus' stops near the entrance of the hospitals and emergency care units (Supplementary Table 1). Seventeen samples were positive for SARS-CoV-2 RNA at public squares distributed across seven different places, corresponding to 34.7% of the positive samples found in this study. Most of the positive samples were detected in benches, with Ct values ranging from 32 to 37.5 (N1) and 34.4 to 39 (N2). We also detected SARS-CoV-2 in different bus terminals in the city, being seven positive samples (14.3%) with Ct values ranging from 29 to 34.7 (N1) and 30.5 to 38.5 (N2). All of these samples were collected at entrance handrails of the bus terminals (Table 1). Finally, we detected the presence of SARS-CoV-2 RNA at very low concentration in one sample collected in the wall of the major public market of the city (N1 Ct = 36.9; N2 Ct = 39.6), and four samples collected at bus stops [floor (1) and benches (3)] in front of public banks (2), sport club (1), and a government department (1), with Ct values ranging from 34.7 to 38.1 (N1) and 37.8 to 39.5 (N2) (Table 1 and Supplementary Table 1). After RNA quantification, we

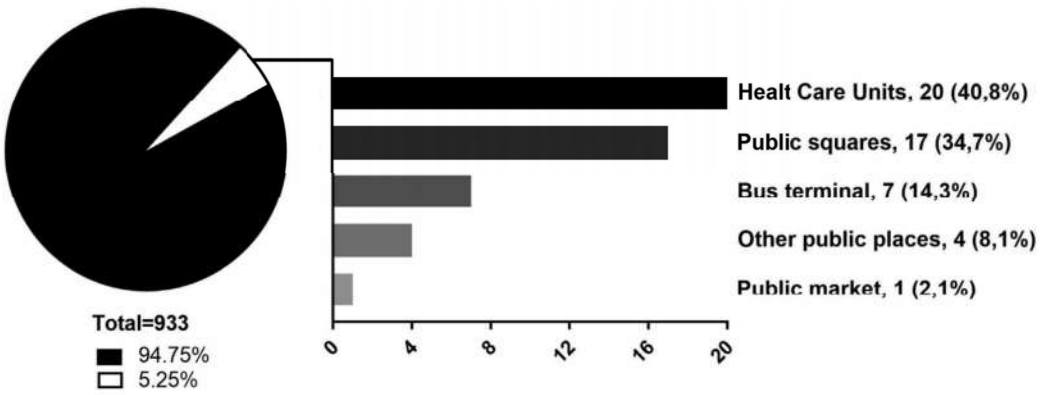


Fig. 1. Distribution of positive samples for SARS-CoV-2 RNA. 933 samples were collected at different locations in Belo Horizonte, Brazil, with 49 being positive for viral genome detection, distributed among five different categories of locations. Raw numbers and percentages are indicated.

observed that most positive samples had very low viral amount (≤1 genomic unit/10 cm^2), and the higher viral load were detected in samples collected from health care units (Table 1).

It is important to mention that samples were not collected evenly among the different categories of places. Most of the samples were collected from health care units (403 samples, 43.2%) and public squares (269 samples, 28.8%) (Supplementary Fig. 2), and it can be the reason why over 3/4 of positive samples were detected from samples collected at these locations. Considering the proportion of positivity in the different places, bus terminals exhibit the higher positivity rate among the evaluated places, followed by public market, public squares, and health care units. The Belo Horizonte City Hall was informed about the contaminated areas, and, after disinfection (laundry detergent followed by 1% sodium hypochlorite), viral RNA could no longer be detected, except for one sample recollected at the entrance of a hospital. We did not identify any positive samples in public parks, education centers, and a mall that were included in this study.

Samples were collected from different materials, the majority being from concrete (383 samples, 41.1%) and metal (282 samples, 30.2%), and also rock, brickwork, wood, glass, and a minority of other materials, including plastic and pottery (10 samples each) and asphalt (4 samples) (Fig. 2A). From the 49 positive samples, 20 were identified in metal surfaces, especially from benches of bus stops and handrails of hospitals entrances and bus terminals (Fig. 2B and Supplementary Table 1). Fourteen positive samples were recovered from concrete surfaces, mainly sidewalks near health care units and public squares (Figs. 2B and 3). Nine positive samples were found in rock surfaces, all being found in public squares but one sampled from a frontal pillar of a hospital (Fig. 3). SARS-CoV-2 RNA was also identified in two samples collected from brickworks (one from health care unit and other from a public square), and in two plastic surfaces from benches of an emergency care unit (Fig. 3 and Supplementary Table 1). Finally, SARS-CoV-2 RNA was detected in one sample from a wood bench located in a public square, and one sample from a glass surface in a bus stop in front of a hospital (Figs. 2B and 3).

4. Discussion

Although we sought to detect the presence of viral RNA, not infectious particles, it is possible that infectious particles were present in these environments, and care must be taken to avoid further contamination and the eventual collapse of the local health system. The infectious dose of SARS-CoV-2 virions to start a productive infection in humans is still unclear, and also if infectious particles can be recovered from surfaces with low viral load. Previous studies have reported the isolation of SARS-CoV-2 from nasopharyngeal and oropharyngeal samples from distinct patients in a nursing facility (USA) with high Ct

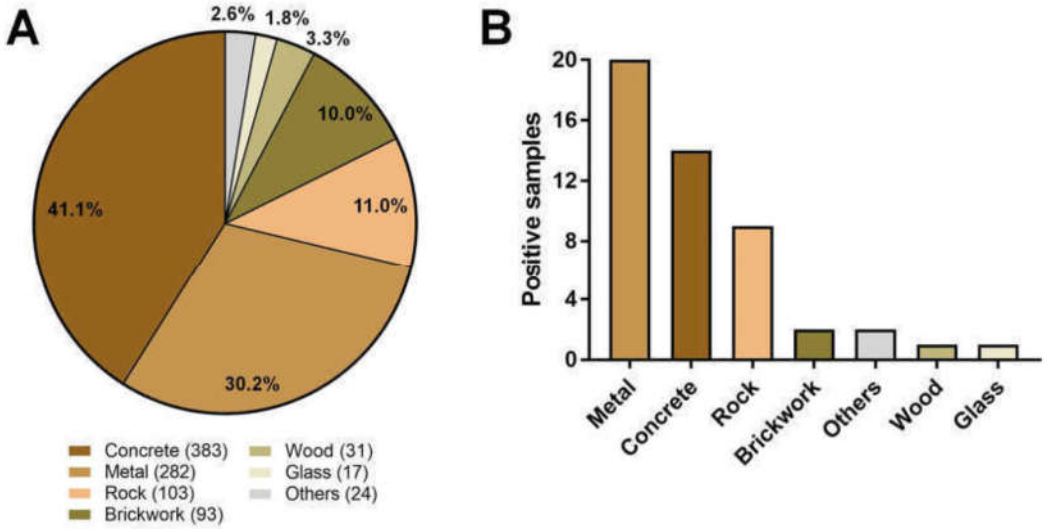


Fig. 2. Environmental samples collected from different materials. (A) Distribution of samples collected in different surfaces. Raw numbers of samples are indicated in front of the materials names; (B) Distribution of positive samples among different kinds of materials.

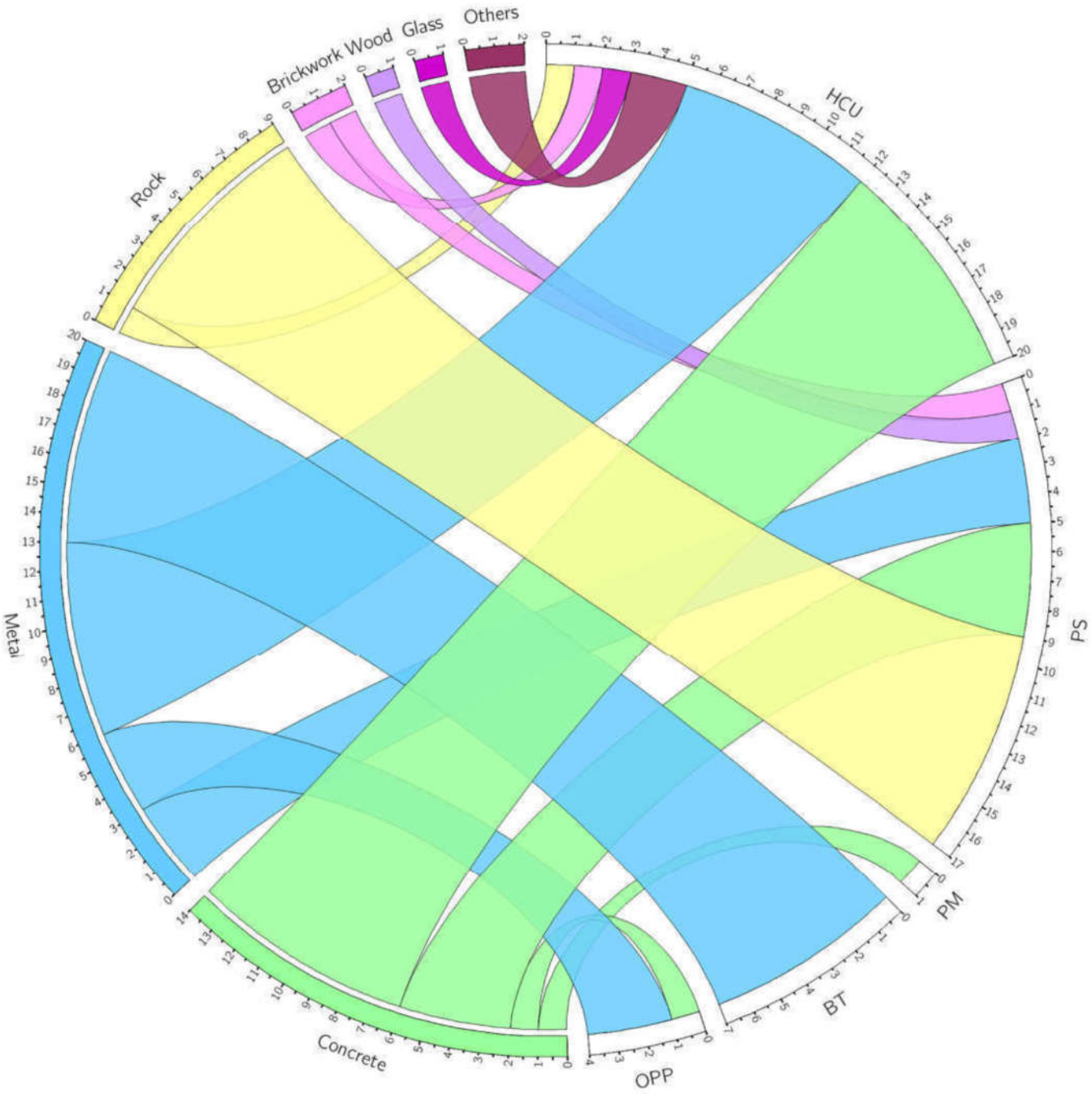


Fig. 3. Association between locations and materials of positive samples for SARS-CoV-2 RNA. Circos plot associating locations and surface materials of positive samples. HCU: health care units; PS: public squares; PM: public market; BT: bus terminals; OPP: other public places.

value (≤34) (Arons et al., 2020). Furthermore, other studies have reported the recovery of the virus from different surfaces, including metal, cardboard, and plastic (Van Doremalen et al., 2020). Different studies have reported the detection of SARS-CoV-2 genome in different environmental surfaces, especially inside hospital facilities, suggesting that environmental contamination by the virus is possible (Carraturo et al., 2020; Chia et al., 2020; Jiang et al., 2020; Wu et al., 2020; Ye et al., 2020). Chin and coworkers observed that SARS-CoV-2 remains viable in smooth surfaces from distinct materials, including metal and plastic, up to 4 days at 22 °C and humidity around 65% (Chin et al., 2020). Therefore, it is possible that infectious particles could also be recovered from the surfaces found positive for the virus RNA in this study, especially considering that around 40% of the positive samples identified in this study had Ct values below 34. Although the real infectious potential of viruses detected in this study cannot be established, we believe the detection of SARS-CoV-2 by molecular assays indicates the potential risk of infection, and care must be taken to avoid further increases in the number of COVID-19 cases. It is important to mention that most of the positive samples in our study had very low amount of RNA (≤1 genomic unit/cm^2) and such amount of virus might not be able to trigger COVID-19 in patients. Nevertheless, the viral RNA in different surfaces was detected in our assays, despite the low concentration, and care must be taken to avoid possible infection.

Although Belo Horizonte has relatively few cases when compared with other cities in Brazil, such as São Paulo and Rio de Janeiro, our data reinforce that the virus is circulating in the city and can be found

Table 1
Positive samples for SARS-CoV-2 RNA in surfaces at different locations in Belo Horizonte, Brazil.

Category	Sample ID	Surface (material)	Ct value (N1)	Concentration (N1) (genomic units/10 cm^2)
Health care units	HCU 29C	Wall (brickwork)	25.9	67.9
	HCU 2 B	Floor (concrete)	34.1	≤1
	HCU 3 N	Floor (concrete)	37.2	≤1
	HCU 8 A	Floor (concrete)	28.5	11.7
	HCU 12 E	Floor (concrete)	37.7	≤1
	HCU 13 J	Wall (concrete)	34.5	≤1
	HCU 29 A	Sidewalk (concrete)	34.2	≤1
	HCU 29 M	Sidewalk (concrete)	36.8	≤1
	HCU 30C	Floor (concrete)	36.2	≤1
	HCU 3 D	Bench (metal)	34.1	≤1
	HCU 3 I	Bench (metal)	34	≤1
	HCU 1 D	Bench (metal)	31.6	≤1
	HCU 9H	Bench (metal)	36.7	≤1
	HCU 11C	Handrail (metal)	36.4	≤1
	HCU 9 A	Bench (metal)	33.1	≤1
	HCU 13 E	Bench (metal)	36.9	≤1
	HCU 21 D	Pillar (rock)	36	≤1
	HCU 29 F	Bench (plastic)	23.7	299.8
	HCU 29H	Bench (plastic)	23.3	392.8
	HCU 12 D	(glass)	36	≤1
Public squares	PS 18 R	Wall (brickwork)	37.4	≤1
	PS 1 E	Table (concrete)	33.1	≤1
	PS 3C	Bench (concrete)	35.8	≤1
	PS 6 G	Bench (concrete)	37.5	≤1
	PS 18 X	Sidewalk (concrete)	35.4	≤1
	PS 15 F	Bench (wood)	34.8	≤1
	PS 3 K	Bench (metal)	35.7	≤1
	PS 17 J	Handrail (metal)	36.5	≤1
	PS 18 I	Handrail (metal)	35.1	≤1
	PS 2 G	Bench (rock)	35.6	≤1
	PS 2H	Bench (rock)	34	≤1
	PS 2 I	Bench (rock)	33.9	≤1
	PS 9C	Bench (rock)	35.5	≤1
	PS 17 A	Bench (rock)	32.3	≤1
	PS 17 D	Bench (rock)	32	≤1
	PS 17 E	Bench (rock)	34	≤1
	PS 18 M	Bench (rock)	36.9	≤1
Bus terminals	BT 1 A	Handrail (metal)	32.8	≤1
	BT 1 B	Handrail (metal)	34.5	≤1
	BT 1C	Handrail (metal)	29	8.4
	BT 1 Q	Handrail (metal)	32.1	≤1
	BT 1 S	Handrail (metal)	32.8	≤1
	BT 1 W	Handrail (metal)	29	8.4
	BT 2 D	Handrail (metal)	34.7	≤1
Other public places	OPP 11C	Floor (concrete)	36.7	≤1
	OPP 5 A	Bench (metal)	36.2	≤1
	OPP 10 A	Bench (metal)	38.1	≤1
	OPP 11 A	Bench (metal)	34.7	≤1
Public market	PM 1 G	Wall (concrete)	36.9	≤1

For PCR quantification, we consider 1 SARS-CoV-2 plaque forming unit as 1 SARS-CoV-2 genomic unit.

on surfaces such as benches, tables, handrails, and floors, and in places with a large flow of people, such as public squares and hospital entrances. The detection of the viral RNA at these sites indicates that adequate cleaning of public environments and reinforcement of educational campaigns for hygienic and social distancing practices should be undertaken. Given the short sampling period in this study, our data do not support a correlation between different environmental parameters, such as local temperature or humidity, and the detection of the virus. Nevertheless, other studies addressed this matter and suggested that SARS-CoV-2 remains viable for up to 7 days at 22 °C, precisely the average temperature in the days we made the sampling for virus detection for this study (Chin et al., 2020). Furthermore, in our work, a mean of 54% of humidity was observed, and another human coronavirus, HCoV-229E, was more stable at 50% of humidity, a similar humidity value (Kampf et al., 2020). These data support the indication that viable viruses could be recovered from places where we detected the viral RNA. Further studies may contribute to elucidate better how frequent we can isolate SARS-CoV-2 from inanimate surfaces in the field.

It is important to notice that the transmission of the virus in buses seems to be considerably high, as demonstrated in China, where in a bus with 68 individuals, one person in a 100 min trip disseminated the virus to another 23 passengers (Shen et al., 2020). Although we did not evaluate the presence of the virus inside the buses of our city, the detection in the bus terminals and stations is an indicative that those places are an important risk factor for people to get infected, not only inside the buses. Our study highlights the need for the constant assessment of the presence of the virus, not only in hospital facilities, but also in places close to medical areas and with a large circulation of people. The presence of SARS-CoV-2 in these environments can result in an increase in the number of cases of the disease in the near future if control measures are not forcefully adopted.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.142645.

CRediT authorship contribution statement

Jônatas Santos Abrahão: Conceptualization, Supervision, Resources, Writing - original draft, Writing - review & editing. **Lívia Sacchetto:** Investigation, Writing - review & editing. **Izabela Maurício Rezende:** Investigation. **Rodrigo Araújo Lima Rodrigues:** Investigation, Visualization,

Writing - original draft, Writing - review & editing. **Ana Paula Correia Crispim:** Investigation. **César Moura:** Investigation. **Diogo Correia Mendonça:** Investigation. **Erik Reis:** Investigation, Writing - review & editing. **Fernanda Souza:** Investigation. **Gabriela Fernanda Garcia Oliveira:** Investigation. **Iago Domingos:** Investigation. **Paulo Victor de Miranda Boratto:** Investigation. **Pedro Henrique Bastos Silva:** Investigation. **Victoria Fulgêncio Queiroz:** Investigation. **Talita Bastos Machado:** Investigation. **Luis Adan Flores Andrade:** Investigation. **Karine Lima Lourenço:** Investigation. **Thaís Silva:** Investigation. **Graziele Pereira Oliveira:** Investigation. **Viviane de Souza Alves:** Conceptualization, Writing - original draft. **Pedro Augusto Alves:** Investigation. **Erna Geessien Kroon:** Conceptualization, Writing - original draft. **Giliane de Souza Trindade:** Conceptualization, Writing - original draft. **Betânia Paiva Drumond:** Conceptualization, Supervision, Writing - original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank our colleagues from Belo Horizonte City Hall, in especial Coronel Genedempsey Bicalho Cruz (Superintendência de Limpeza Urbana – SLU/BH). In addition, we thank the Gabinete da Reitora da UFMG, the Pró Reitoria de Pesquisa da UFMG/Secretaria de Educação Superior/Ministério da Educação (number 23072.211119/2020-10), Finep/RTR/PRPq/Rede COVID-19 (number 0494/20-0120002600), Programa de Pós-graduação em Microbiologia da UFMG, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPEMIG (Fundação de Amparo à Pesquisa do estado de Minas Gerais) for their financial support. B.P.D, E.G.K., G.S.T. and J.S.A. are CNPq researchers. J.S.A. and P.A.A. are members of Rede Vírus MCTIC.

References

Arons, M.M., Hatfield, K.M., Reddy, S.C., Kimball, A., James, A., Jacobs, J.R., Taylor, J., Spicer, K., Bardossy, A.C., Oakley, L.P., Tanwar, S., Dyal, J.W., Harney, J., Chisty, Z., Bell, M., Methner, M., Paul, P., Carlson, C.M., McLaughlin, H.P., Thornburg, N., Tong, S., Tamin, A., Tao, Y., Uehara, A., Harcourt, J., Clark, S., Brostrom-Smith, C., Page, L.C., Kay, M., Lewis, J., Montgomery, P., Stone, N.D., Clark, T.A., Honein, M.A., Duchin, J.S., Jernigan, J.A., 2020. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. N. Engl. J. Med. 382, 2081–2090. https://doi.org/10.1056/NEJMoa2008457.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622. https://doi.org/10.1373/clinchem.2008.112797.

Carraturo, F., Del Giudice, C., Morelli, M., Cerullo, V., Libralato, G., Galdiero, E., Guida, M., 2020. Persistence of SARS-CoV-2 in the environment and COVID-19 transmission risk from environmental matrices and surfaces. Environ. Pollut. https://doi.org/10.1016/j.envpol.2020.115010.

Centers for Disease Control and Prevention Division of Viral Diseases, 2020. CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel.

Chia, P.Y., Coleman, K.K., Tan, Y.K., Ong, S.W.X., Gum, M., Lau, S.K., Lim, X.F., Lim, A.S., Sutjipto, S., Lee, P.H., Son, T.T., Young, B.E., Milton, D.K., Gray, G.C., Schuster, S., Barkham, T., De, P.P., Vasoo, S., Chan, M., Ang, B.S.P., Tan, B.H., Leo, Y.S., Ng, O.T., Wong, M.S.Y., Marimuthu, K., Lye, D.C., Lim, P.L., Lee, C.C., Ling, L.M., Lee, L., Lee, T.H., Wong, C.S., Sadarangani, S., Lin, R.J., Ng, D.H.L., Sadasiv, M., Yeo, T.W., Choy, C.Y., Tan, G.S.E., Dimatatac, F., Santos, I.F., Go, C.J., Chan, Y.K., Tay, J.Y., Tan, J.Y.L., Pandit, N., Ho, B.C.H., Mendis, S., Chen, Y.Y.C., Abdad, M.Y., Moses, D., 2020. Detection of air and surface contamination by SARS-CoV-2 in hospital rooms of infected patients. Nat. Commun. https://doi.org/10.1038/s41467-020-16670-2.

Chin, A.W.H., Chu, J.T.S., Perera, M.R.A., Hui, K.P.Y., Yen, H.-L., Chan, M.C.W., Peiris, M., Poon, L.L.M., 2020. Stability of SARS-CoV-2 in different environmental conditions. The Lancet Microbe. https://doi.org/10.1016/s2666-5247(20)30003-3.

Cutler, D., 2020. How will COVID-19 affect the health care economy? JAMA - J. Am. Med. Assoc. https://doi.org/10.1001/jama.2020.7308.

de Almeida, P.R., Demoliner, M., Antunes Eisen, A.K., Heldt, F.H., Hansen, A.W., Schallenberger, K., Fleck, J.D., Spilki, F.R., 2020. SARS-CoV2 quantification using RT-dPCR: a faster and safer alternative to assist viral genomic copies assessment using RT-qPCR. bioRxiv https://doi.org/10.1101/2020.05.01.072728.

Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J., 2020. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. https://doi.org/10.1038/s41564-020-0695-z.

Guo, Z.D., Wang, Z.Y., Zhang, S.F., Li, X., Li, L., Li, C., Cui, Y., Fu, R. Bin, Dong, Y.Z., Chi, X.Y., Zhang, M.Y., Liu, K., Liu, K., Cao, C., Liu, B., Zhang, K., Gao, Y.W., Lu, B., Chen, W., 2020. Aerosol and surface distribution of severe acute respiratory syndrome coronavirus 2 in hospital wards, Wuhan, China, 2020. Emerg. Infect. Dis. 26, 1583–1591. https://doi.org/10.3201/eid2607.200885.

INMET, 2020. Normas Climatológicas do Brasil. URL. https://portal.inmet.gov.br/.

Jiang, F.C., Jiang, X.L., Wang, Z.G., Meng, Z.H., Shao, S.F., Anderson, B.D., Ma, M.J., 2020. Detection of severe acute respiratory syndrome coronavirus 2 RNA on surfaces in quarantine rooms. Emerg. Infect. Dis. https://doi.org/10.3201/eid2609.201435.

Kampf, G., Todt, D., Pfaender, S., Steinmann, E., 2020. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. J. Hosp. Infect. https://doi.org/10.1016/j.jhin.2020.01.022.

Lai, J., Ma, S., Wang, Y., Cai, Z., Hu, J., Wei, N., Wu, J., Du, H., Chen, T., Li, R., Tan, H., Kang, L., Yao, L., Huang, M., Wang, H., Wang, G., Liu, Z., Hu, S., 2020. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw. Open https://doi.org/10.1001/jamanetworkopen.2020.3976.

Liu, Yuan, Ning, Z., Chen, Y., Guo, M., Liu, Yingle, Gali, N.K., Sun, L., Duan, Y., Cai, J., Westerdahl, D., Liu, X., Xu, K., Ho, K. fai, Kan, H., Fu, Q., Lan, K., 2020. Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. Nature. https://doi.org/10.1038/s41586-020-2271-3.

Secretaria Municipal de Saúde, 2020. Boletim epidemiológico e assistencial - COVID-19 N°108/2020 (Belo Horizonte).

Shen, Y., Li, C., Dong, H., Wang, Z., Martinez, L., Sun, Z., Handel, A., Chen, Z., Chen, E., Ebell, M.H., Wang, F., Yi, B., Wang, H., Wang, X., Wang, A., Chen, B., Qi, Y., Liang, L., Li, Y., Ling, F., Chen, J., Xu, G., 2020. Community outbreak investigation of SARS-CoV-2 transmission among bus riders in eastern China. JAMA Intern. Med. https://doi.org/10.1001/jamainternmed.2020.5225.

Van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., Lloyd-Smith, J.O., De Wit, E., Munster, V.J., 2020. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N. Engl. J. Med. https://doi.org/10.1056/NEJMc2004973.

Wang, J., Feng, H., Zhang, S., Ni, Z., Ni, L., Chen, Y., Zhuo, L., Zhong, Z., Qu, T., 2020. SARS-CoV-2 RNA detection of hospital isolation wards hygiene monitoring during the Coronavirus Disease 2019 outbreak in a Chinese hospital. Int. J. Infect. Dis. https://doi.org/10.1016/j.ijid.2020.04.024.

WHO, 2020. Coronavirus disease (COVID) - Weekly Epidemiological and Operational updates September 2020–21 September 2020.

Wu, S., Wang, Y., Jin, X., Tian, J., Liu, J., Mao, Y., 2020. Environmental contamination by SARS-CoV-2 in a designated hospital for coronavirus disease 2019. Am. J. Infect. Control https://doi.org/10.1016/j.ajic.2020.05.003.

Ye, G., Lin, H., Chen, S., Wang, S., Zeng, Z., Wang, W., Zhang, S., Rebmann, T., Li, Y., Pan, Z., Yang, Z., Wang, Y., Wang, F., Qian, Z., Wang, X., 2020. Environmental contamination of SARS-CoV-2 in healthcare premises. J. Inf. Secur. https://doi.org/10.1016/j.jinf.2020.04.034.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2001017.

7. Conclusões gerais:

Estudos de prospecção de vírus gigantes têm revelado nos últimos anos um universo notável de diversidade viral em nosso planeta. Alguns representantes, como os mimivírus, parecem ser mais abundantes e ubíquos, contendo centenas de isolados já relatados. Já outras famílias virais parecem ter no ambiente uma representação bastante rara, assim como observado para as poucas sequências de yaravírus detectadas em bancos de dados de metagenômica de diversos ambientes ao redor do mundo. O yaravírus descrito nesse trabalho representa o primeiro isolado viral que infecta culturas de *Acanthamoeba spp.* e que potencialmente pode não pertencer aos membros do filo *Nucleocytoviricota,* diferindo fortemente de todos os outros vírus de ameba já descritos até o momento. Este vírus possui um genoma de dupla fita de DNA com um número notável de ORFans (mais de 90%), incluindo novas proteínas estruturais. Devido à falta de informações filogenéticas, suas origens ainda são um mistério e a análise de similaridade com outros vírus é bastante complicada. As novas características observadas no yaravírus nos levaram a propor a criação de um novo gênero ("Yaravirus") e família viral ("Yaraviridae"), tendo a espécie "Yaravirus brasiliense" como primeiro membro. Um outro representante dos grupos de vírus gigantes que apresenta uma limitada quantidade de isolados são os pandoravírus. Por causa dessa limitação, muitos estudos consideram o pangenoma desses organismos ainda em aberto, havendo muitas lacunas quanto à biologia desses vírus gigantes. Nesse trabalho, além das etapas básicas do ciclo de multiplicação de diferentes isolados de pandoravírus, pudemos também identificar diversos estágios de modificação estrutural pelos quais as células hospedeiras são submetidas durante o curso da infecção. Apesar disso, é importante ressaltar que muitos aspectos desse ciclo ainda permanecem em aberto, com este trabalho fornecendo novos dados quanto ao curso da infecção e também revelando novas questões que estudos futuros ainda precisam responder, especialmente ao nível molecular.

Por fim, um outro questionamento aberto pelos estudos de prospecção e caracterização de vírus gigantes foi o relacionado ao processo de formação dessas partículas virais que muitas vezes se apresentam compostas por quantidades que podem chegar a centenas de proteínas. Entender os componentes envolvidos na formação de um vírion acaba por contribuir para a discussão sobre questões relacionados com a evolução dos vírus, suas origens e sobre as etapas iniciais de seu ciclo. As análises filogenéticas das 127 proteínas de tupanvírus indicam que uma porção substancial da estrutura da partícula foi obtida por eventos de transferência gênica a partir de organismos celulares, pertencentes aos três domínios da vida. No entanto, ainda é

considerado um campo aberto de estudos saber se essa porção de genes provenientes de outros domínios seria algo relevante quando comparado com outros membros do NCLDV. O estudo do proteoma de outros vírus gigantes se torna dessa forma um campo de estudo bastante interessante, assim também como a análise dos dados filogenéticos dos vírus que venham a ter sua proteômica determinada, buscando entender qual a contribuição de outros domínios para a formação de sua estrutura.

8. Produção científica e outras atividades

Doutorado sanduíche, pelo período de 1 ano, na *Aix-Marseille Université*. Supervisor estrangeiro: Bernard La Scola

Primeira autoria no artigo de revisão *"The multiple origins of proteins present in tupanvirus particles"*, publicado no periódico *Current Opinion in Virology*.

Primeira autoria do artigo "*New Isolates of Pandoraviruses: Contribution to the Study of Replication Cycle Steps",* publicado no periódico *Journal of Virology.*

Primeira autoria no artigo *"A novel 80-nm virus infecting Acanthamoeba castellanii"*, publicado no periódico *Proceedings of the National Academy of Sciences.* Esse artigo também se encontra disponível na forma de *pre-print* no periódico *BioRxiv*. Essa versão alcançou a maior popularidade dentre as publicações de mesma época presentes no periódico em menos de 15 dias da sua aceitação.

Primeira autoria no artigo "*Yaraviridae": a new family of virus infecting Acanthamoeba castellanii*", publicado no periódico *Archives of Virology.*

Primeira autoria no artigo de revisão *"Giants, rise: a brief history of giant viruses' studies in Brazil"*, publicado no periódico *Viruses*.

Co-autoria no artigo *"Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere"*, publicado no periódico *Nature Communications.* (Esse artigo foi considerado entre os 50 mais importantes publicados pela revista em 2018).

Co-autoria no artigo *"Single Cell Micro-aspiration as an Alternative Strategy to Fluorescence-activated Cell Sorting for Giant Virus Mixture Separation"*, publicado no periódico *Journal of Visualized Experiments*.

Co- autoria no artigo "*Detection of SARS-CoV-2 RNA on public surfaces in a densely populated urban area of Brazil"*, publicado no periódico *Science of The Total Environment.*

Palestrante convidado da disciplina "Seminários em Bioinformática", em uma palestra da série BrazWebinars, abordando o tema "*A mysterious 80 nm amoeba virus with a near-complete ORFan genome challenges the classification of DNA viruses*".

Palestrante convidado no I Ciclo de Palestras Online da Biologia da Universidade Estadual do Piauí/Teresina abordando a palestra de tema "O universo em expansão dos vírus gigantes de amebas: caracterizações biológicas e genéticas dos isolados".

Palestrante convidado no WEB Mini Curso sobre vírus gigantes estabelecido durante a I Mini Reunião Anual Virtual da Sociedade Brasileira pelo Progresso da Ciência.

Professor convidado na aula "Antivirais" para a disciplina de "Virologia Básica e Técnicas Clássicas e Moleculares em Virologia" no curso de pós-graduação em Ciências da Saúde do Instituto René Rachou – Fiocruz Minas.

Trabalho autônomo na na empresa Simile Medicina Diagnóstica auxiliando no diagnóstico de pacientes com suspeita de COVID-19.

Trabalho de extensão auxiliando no diagnóstico de pacientes com suspeita de COVID-19 em uma parceria da Universidade Federal de Minas Gerais com a Secretária de Saúde de Minas Gerais.