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

TESE DE DOUTORADO

Isolamento e Caracterização de Vírus Gigantes em Biomas do Brasil e Antártica

ANA CLÁUDIA DOS SANTOS PEREIRA ANDRADE

Belo Horizonte
2019

ANA CLÁUDIA DOS SANTOS PEREIRA ANDRADE

Isolamento e Caracterização de Vírus Gigantes em Biomas do Brasil e Antártica

Tese de doutorado apresentado ao Programa de Pós-Graduação em Microbiologia da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de **Doutora em microbiologia.**

Orientador: Prof. Jônatas Santos Abrahão

Coorientador: Prof. Danilo Bretas de Oliveira

Belo Horizonte

2019

043 Andrade, Ana Cláudia dos Santos Pereira.
Isolamento e caracterização de vírus gigantes em biomas do Brasil e Antártica [manuscrito] / Ana Cláudia dos Santos Pereira Andrade. – 2019.

217 f. : il. ; 29,5 cm.

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

1. Microbiologia. 2. Vírus Gigantes. 3. Mimiviridae. 4. Prospecção - Brasil. I. Abrahão, Jônatas Santos. II. Oliveira, Danilo Bretas de. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 579

Ficha Catalográfica elaborada pela Biblioteca do Instituto de Ciências Biológicas

DEDICATÓRIA

Aos meus pais, Maria Célia e Célio, às minhas irmãs e ao meu marido Raffael, pelo apoio incondicional desde o início dessa jornada, sem o qual eu não teria chegado até aqui

AGRADECIMENTOS

À Deus, que esteve comigo em todas as horas que precisei;

Ao meu marido e a minha família pelo apoio incondicional e por todo o amor que me transmitem. Se aqui estou concluindo mais uma etapa da minha vida profissional, é graças a vocês.

Ao professor Jônatas Santos Abrahão, pela orientação impecável, por ter me dado a oportunidade de desenvolver esse trabalho, por sempre propor novos desafios e por acreditar no meu potencial. Ao professor Danilo pela coorientação e pelos conselhos;

Aos professores e colegas do Laboratório de Vírus da UFMG, por propiciarem um ambiente agradável de trabalho e um aprendizado tão rico e fundamental para minha formação;

Aos amigos do GEPVIG, pela verdadeira parceria ao longo destes anos. Vocês têm o meu profundo respeito e admiração.

Ao Centro de Microscopia da UFMG e seus ótimos profissionais, pelo excelente trabalho prestado que tanto enriqueceram esse trabalho;

À professora Juliana Cortines pela confiança em meu trabalho e por ter me fornecido as amostras coletadas na Antártica;

Ao Ministério da Saúde pela parceria e pelo fornecimento das amostras clínicas analisadas neste trabalho;

Aos membros da banca avaliadora, pelo interesse, disponibilidade e dedicação para colaborar com a minha formação nesta fase final do trabalho;

As agências financiadoras, CNPq, CAPES e FAPEMIG, pelo auxílio financeiro que permitiu a realização desse trabalho;

A todos vocês, meus sinceros agradecimentos!

“Quando nasci um anjo esbelto,
desses que tocam trombeta, anunciou:
vai carregar bandeira.
Cargo muito pesado pra mulher,
esta espécie ainda envergonhada.
Aceito os subterfúgios que me cabem,
sem precisar mentir...
Mulher é desdobrável. Eu sou.”

(Adélia Prado, 1993)

RESUMO

Tese de Doutorado
Programa de Pós-Graduação em Microbiologia
Universidade Federal de Minas Gerais

ISOLAMENTO E CARACTERIZAÇÃO DE VÍRUS GIGANTES EM BIOMAS DO BRASIL E ANTÁRTICA

ANA CLÁUDIA DOS SANTOS PEREIRA ANDRADE
Orientador: Prof. Jônatas Santos Abrahão
Belo Horizonte, Março de 2019

Os vírus gigantes são conhecidos por serem vírus de morfologia e genomas complexos e nos últimos anos vários novos representantes têm sido descritos. Entretanto, muitos aspectos acerca da diversidade, distribuição e interação destes vírus com as células hospedeiras ainda permanecem desconhecidos. Neste trabalho, descrevemos a prospecção de vírus gigantes usando diversas amostras do Brasil e da Antártica. No total, foram isolados 68 vírus, que foram identificados por PCR em tempo real, microscopia óptica e eletrônica (2 marseillevirus, 3 pandoravírus, 1 cedratvírus e 64 mimivírus). Além disso, analisamos algumas etapas do ciclo de multiplicação de dois grupos virais isolados: os mimivírus e os pandoravírus. Foi demonstrado que os mimivírus dependem da fagocitose e da acidificação dos fagossomos para sua penetração e desnudamento, respectivamente, em hospedeiros amebianos. Os capsídeos parecem ser formados a partir de estruturas lamelares crescentes e adquirem as fibrilas em uma área específica ao redor da fábrica viral (FV), sendo que a aquisição de fibrilas e genoma pela partícula pode ocorrer simultaneamente. Foi observado também que partículas defectivas podem ser formadas mesmo na ausência de virófagos. O estudo do ciclo de multiplicação dos pandoravírus mostrou uma intensa reorganização citoplasmática com formação de uma grande FV apresentando intenso recrutamento de membrana. Nas FV ocorre a morfogênese, e esta pode ser iniciada pelas duas extremidades da partícula. Nas etapas mais avançadas do ciclo, o núcleo não é mais observado e a progênie formada pode ser liberada por exocitose e por lise celular. Em conjunto, estes resultados contribuem para a construção do nosso entendimento a respeito da diversidade e ciclo de multiplicação de alguns vírus gigantes.

Palavras-chave: Vírus gigantes, Prospecção, Brasil, Pandoravírus, Mimivírus

ABSTRACT
Doctoral Thesis
Programa de Pós-Graduação em Microbiologia
Universidade Federal de Minas Gerais

ISOLATION AND CHARACTERIZATION OF GIANT VIRUSES IN BIOMAS OF BRAZIL AND ANTARCTICA

ANA CLÁUDIA DOS SANTOS PEREIRA ANDRADE
Orientador: Prof. Jônatas Santos Abrahão
Belo Horizonte, Março de 2019

Giant viruses are complex members of the virosphere and in recent years several new representatives have been discovered. However, many aspects about the diversity, distribution, and interaction of these viruses with host cells remain unknown. In this work, we describe the prospection of giant viruses using several samples from Brazil and Antarctica. In total, 68 viruses were isolated which were identified by real-time PCR, electron and light microscopy (1 marseillevirus, 2 pandoravirus, 1 cedratvirus and 64 mimivirus). In addition, we have analyzed some stages of the replication cycle of two isolated viral groups: mimivirus and pandoravirus. We demonstrated that mimiviruses depend on phagocytosis and phagosome acidification for entry and uncoating step, respectively, in amoebae hosts. The mimivirus capsids morphogenesis seems to be assembled from growing lamellar structures, the genome and fibrils can be acquired simultaneously, and there is a specific area surrounding the viral factory where particles acquire the surface fibrils. Finally, defective particles can be formed even in the absence of virophages. The study of the replication cycle of pandoraviruses showed an intense cytoplasmic reorganization with formation of a large VF presenting intense recruitment of membrane. In the VF morphogenesis occurs, and this can be initiated by both ends of the particle. In the more advanced stages of the cycle, the nucleus is no longer observed and the progeny formed can be released by exocytosis and by cell lysis. Together, these results contribute to building our understanding of the diversity and replication cycle of some giant viruses.

Key-words: Giant viruses, Prospecting, Brazil, Pandoravirus, Mimivirus

SUMÁRIO

1. INTRODUÇÃO 9
2. JUSTIFICATIVA 18
3. OBJETIVOS 19
3.1 Objetivo geral: 19
3.2 Objetivos específicos: 19
4. METODOLOGIA, RESULTADOS E DISCUSSÕES PARCIAIS 20
4.1. Artigo # 1: Ubiquitous Giants: A plethora of giant viruses found in Brazil and Antarctica 20
4.2. Artigo #2: Filling gaps about mimivirus entry, uncoating and morphogenesis 36
4.3. Artigo #3: New isolates of pandoraviruses: contribution to the study of replication cycle steps 52
5. DISCUSSÃO 66
6. CONCLUSÕES 74
7. REFERÊNCIAS BIBLIOGRÁFICAS 77
8. OUTRAS ATIVIDADES DESENVOLVIDAS DURANTE O DOUTORADO 87
8.1 Participação de eventos científicos 87
8.2 Supervisão de alunos de iniciação científica 87
8.3 Artigos completos publicados em periódicos indexados durante o período do doutorado (Anexos) 87
8.4 Trabalhos aceitos para publicação 89
8.5 Disciplinas cursadas e aproveitamento 89

1. INTRODUÇÃO

A descrição do primeiro vírus gigante associado à ameba, denominado *Acanthamoeba polyphaga mimivirus* (APMV), em 2003, representa um marco na virologia e revolucionou conceitos clássicos acerca da classificação e definição dos vírus (LA SCOLA, *et al.*, 2003). Inicialmente, o APMV despertou grande interesse da comunidade científica por apresentar uma partícula estruturalmente complexa com 750 nm de diâmetro, sendo o maior vírus relatado até então. O capsídeo deste vírus possui simetria pseudo-icosaédrica e é formado por uma tripla camada proteica (XIAO *et al.*, 2009). As partículas deste vírus apresentam uma face em formato de estrela denominada de *star-gate*, estrutura responsável pela liberação do genoma viral na célula hospedeira durante o ciclo de multiplicação viral (ZAUBERMAN *et al.,* 2008). Na região interna ao capsídeo, o APMV possui uma membrana lipídica envolvendo o genoma que parece estar imerso em uma matriz fibrosa (revisado por COLSON *et al.*, 2017; XIAO *et al.*, 2009). Além disso, a superfície viral é coberta por uma camada de fibrilas glicoproteicas com 125 nm de comprimento, que são importantes na adesão viral a diferentes organismos, incluindo as amebas (Figura 1A) (RODRIGUES *et al.*, 2015).

Além da complexidade estrutural da partícula, o genoma do APMV também se mostrou surpreendente (RAOULT *et al*., 2004). Este genoma é formado por uma dupla fita de DNA linear composta por 1.2 mega pares de bases (pb) que codificam 1018 genes preditos (LEGENDRE et al., 2011). Alguns destes genes que compõe o arsenal genético dos mimivírus eram considerados como exclusivos de organismos celulares, entre eles destacam-se genes relacionados ao processo transcricional e traducional, como aminoacil-tRNA-sintetases, além de fatores raramente encontrados na virosfera, como RNA transportadores e fatores de tradução (ABRAHÃO *et al.*, 2017; JEUDY *et al.*, 2012; RAOULT *et al.*, 2004). De acordo com o Comitê Internacional de Taxonomia Viral (ICTV), a família *Mimiviridae* é dividida em dois gêneros: *Mimivirus* e *Cafeteriavirus* (FISCHER *et al.*, 2010). Baseado em análises filogenéticas, o gênero *Mimivirus* é dividido em linhagens A, B e C, sendo o APMV protótipo do gênero e representante da linhagem A (ARSLAN *et al.*, 2011; LA SCOLA, B. *et al.*, 2003; YOOSUF *et al.*, 2012).

O espectro de hospedeiros destes vírus geralmente é restrito a poucas espécies de amebas dos gêneros *Acanthamoeba* e *Vermoameba* (ABRAHÃO *et al.*, 2018, LA SCOLA *et al*., 2003, PHILIPPE *et al.,* 2013, LEGENDRE *et al.,* 2015, RETENO *et al.,* 2015, BAJRAI *et al.,* 2016, ANDREANI *et al.,* 2017). Em 2010 outro hospedeiro, o *Cafeteria roenbergensis,* foi descrito para os mimivírus e este vírus foi denominado cafeteria roenbergensis vírus, sendo representante do gênero *Cafeteriavirus* (FISCHER et al., 2010). Descobertas recentes de novos vírus propostos como novos membros da família *Mimiviridae* mostrou que o espectro de hospedeiros pode ser mais amplo, já que estes novos mimivírus são capazes de infectar outras espécies de protozoários como *Bodo saltans* e espécies de peixes, como *Acipenser fulvescens* (CLOUTHIER *et al.,* 2018; DEEG *et al.,* 2018).

Nos anos seguintes, a busca por novos vírus associados a amebas foi intensificada e durante os dezesseis anos de estudo e prospecção de vírus gigantes, centenas de novos isolados foram obtidos de diversas regiões do planeta (ANDREANI *et al.*, 2016, 2017; BAJRAI *et al.*, 2016; BOYER *et al.*, 2009; LEGENDRE *et al.*, 2014; LEGENDRE *et al.*, 2015; PHILIPPE *et al.*, 2013; RETENO *et al.*, 2015). No ano de 2009, um novo vírus foi isolado em cultura de *Acanthamoeba polyphaga* e denominado *Marseillevirus marseillevirus* (MsV) (BOYER *et al.*, 2009). O MsV possui um capsídeo de simetria icosaédrica de 250 nm e um genoma composto por aproximadamente 370 Mb. O isolamento do MsV levou a criação de uma nova família viral denominada *Marseilleviridae*, que foi reconhecida pelo ICTV no ano de 2013 (Figura 1B) (COLSON *et al.*, 2013a).

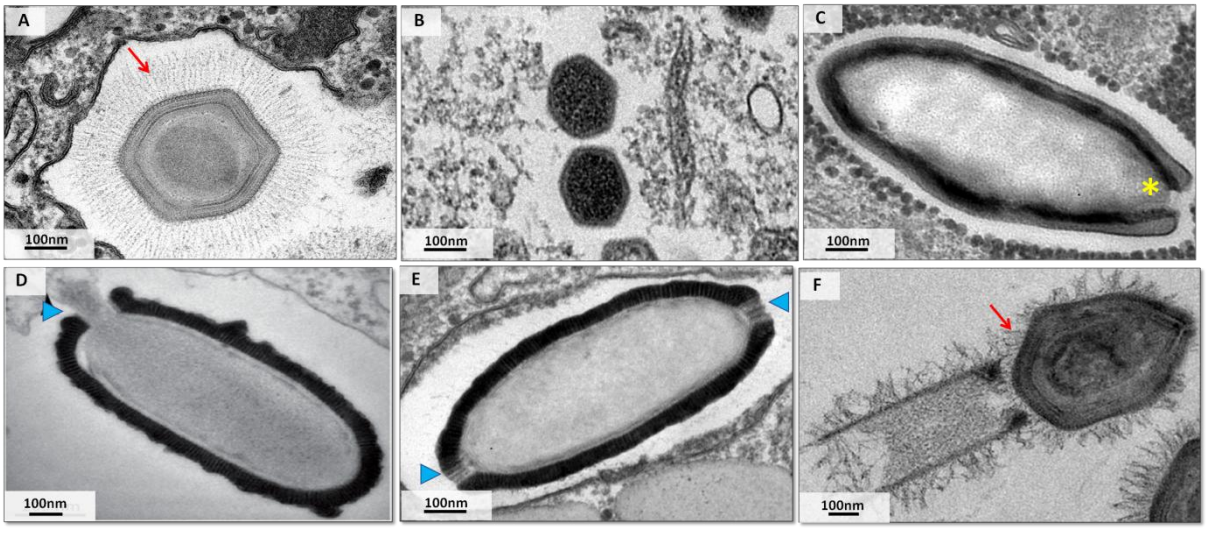


Figura 1: Estrutura de alguns vírus associados a amebas. Imagens de microscopia eletrônica de transmissão, evidenciando as diferenças morfológicas entre os vírus de diferentes grupos. A) Mimivírus B) Marseillevírus; C) Pandoravírus; D) Pithovírus E) Cedratvírus F) Tupanvírus. As setas vermelhas indicam as fibrilas de superfície; o asterisco amarelo indica o ostíolo apical da partícula; a seta azul indica o poro apical da partícula. Fontes: Banco de imagens do Gepvig (2018); Legendre et al., 2014.

No ano de 2013, outro grupo de vírus gigantes foi descoberto quando dois isolados foram obtidos em culturas de *A. castellanii* a partir de amostras de sedimentos de um rio localizado na costa central do Chile e lama de uma lagoa próximo de Melbourne, na Austrália. Estes vírus foram nomeados de pandoravirus salinus e pandoravirus dulcis, respectivamente (PHILIPPE *et al.*, 2013). As partículas de pandoravírus são ovais e possuem um ostíolo apical em uma das extremidades da partícula que apresenta 1 µm de comprimento e 0,5 µm de diâmetro (Figura 1C) (PHILIPPE *et al.*, 2013). O genoma destes vírus é composto por uma dupla fita de DNA linear de até 2.5 Mb (pandoravirus salinus), atualmente o maior genoma conhecido da virosfera. Além disso, esses vírus possuem características estruturais e genômicas distintas aos vírus gigantes já descritos, fazendo com que não fossem alocados em nenhuma família já estabelecida (PHILIPPE *et al.*, 2013). Posteriormente a descoberta destes vírus, outros dez pandoravírus foram descritos em diferentes ambientes de várias partes do mundo, incluindo o Brasil (SCHEID *et al.*,2008; ANTWERPEN *et al.*, 2015; DORNAS *et al.*, 2015; LEGENDRE et al., 2018; AHERFI et al., 2018; LEGENDRE et al., 2019).

Os pithovírus compõem o quarto grupo de vírus gigantes descritos, sendo o primeiro deles isolado em culturas de *A. castellanii* a partir de uma amostra de solo siberiano congelado (conhecido como *permafrost* siberiano) datada de aproximadamente 30.000 anos, e foi denominado de pithovirus sibericum (LEGENDRE *et al.*, 2014). As partículas deste vírus apresentam 1,5 µm de comprimento (Figura 1D) (LEGENDRE *et al*., 2014). Morfologicamente os pithovírus são semelhantes aos pandoravírus, porém, possuem uma estrutura semelhante a uma grade hexagonal ao final do poro apical que é ausente nos pandoravírus (LEGENDRE *et al*., 2014). Em 2016, um vírus com morfologia semelhante ao do pithovírus foi descrito, sendo denominado de cedratvírus e

possui o tamanho de aproximadamente 1,2 µm (ANDREANI *et al.*, 2016). Assim como o pithovírus, o cedratvírus possui poros apicais, porém nas duas extremidades da partícula viral (Figura 1E) (ANDREANI et al., 2016).

As características biológicas compartilhadas pelos vírus gigantes permitiram a sua incorporação em um suposto grupo monofilético denominado Vírus Grandes Núcleo-citoplasmáticos de DNA (NCLDV). Atualmente o grupo NCLDV é composto por sete famílias que infectam hospedeiros variados, incluindo metazoários (*Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*), algas verdes e marrons (*Phycodnaviridae*), e protistas (*Mimiviridae* e *Marseilleviridae*) (YUTIN; WOLF; KOONIN, 2014). Além disso, alguns outros vírus associado a protistas, recentemente descobertos e que ainda não possuem família assentida pelo ICTV fazem parte deste grupo, sendo eles: pandoravírus, faustovírus, cedratvírus, pithovírus, mollivírus, kaumoebavírus, pacmanvírus, orpheovírus e medusavírus (Tabela 1) (ANDREANI *et al.*, 2016, 2017; BAJRAI *et al.*, 2016; LEGENDRE *et al.*, 2014; LEGENDRE *et al.*, 2015; PHILIPPE *et al.*, 2013; RETENO *et al.*, 2015; SCHULZ *et al.*, 2017; ANDREANI et al., 2018; YOSHIKAWA *et al*., 2019).

Os estudos de prospecção mais recentes mostraram que os vírus gigantes são ubíquos, assim como seus hospedeiros. Estes vírus já foram isolados e/ou detectados em todos os continentes da Terra a partir de diversos tipos de amostras como água, esgoto, solo, invertebrados e mamíferos (Tabela 1) (revisado por AHERFI *et al.*, 2016; BOUGHALMI *et al.*, 2013; COLSON *et al.*, 2013; DORNAS *et al.*, 2015; DORNAS *et al.*, 2014; KEREPESI; GROLMUSZ, 2016, 2017; LEGENDRE *et al.*, 2015; PHILIPPE *et al.*, 2013; POPGEORGIEV 2013; ABRAHÃO *et al.*, 2018). Estudos de metagenômica reforçam a ideia de ubiquidade além de indicar que muitos vírus associados a amebas ainda não foram isolados (ZHANG et al., 2015; MIHARA *et al.*, 2018; SCHULZ *et al.,* 2017; SCHULZ *et al.,* 2018). Entretanto, a diversidade, distribuição, assim como o papel destes vírus no ecossistema ainda são pouco compreendidos.

Tabela 1: Principais representantes dos grupos virais associados a amebas

Família/ Grupos	Espécie/isolado representante	Amostra/Local de origem	Célula de isolamento	Morfologia/ Tamanho	Genoma (Kb)	Referência
Mimiviridae*	*Acanthamoeba polyphaga mimivirus**	Água de torre de resfriamento/ Inglaterra	*A. polyphaga*	Pseudo-icosaédrica/ 750 nm	1182	La Scola et al 2003
Marseilleviridae*	*Marseillevirus marseillevirus**	Água de torre de resfriamento/ França	*A.polyphaga*	Icosaédrica/ 250 nm	368	Boyer et al 2009
Pandoravirus	Pandoravirus salinus	Sedimento marinho/ Chile	*A. castellanii*	Ovóide/ 1 µm	2474	Philippe et al 2013
Faustovirus	Faustovirus E12	Esgoto/ França	*V. vermiformes*	Icosaédrico/ 250 nm	466	Reteno et al 2015
Mollivirus	Mollivirus sibericum	Permafrost/ Rússia	*A. castellanii*	Esférico/ 600 nm	652	Legendre et al 2015
Pithovirus	Pithovirus sibericum	Permafrost / Rússia	*A. castellanii*	Ovóide/ 1,5 µm	610	Legendre et al 2014
Cedratvirus	Cedratvirus A11	?/Argélia	*A. castellanii*	Ovóide/ 1,2 µm	589	Andreani et al 2016
Pacmanvirus	Pacmanvirus A23	?/Argélia	*A. castellanii*	Icosaédrico/ 175 nm	395	Andreani et al 2017
Kaumoebavirus	Kaumoebavirus	Esgoto/ Arábia Saudita	*V. vermiformes*	Icosaédrico/ 250 nm	350	Bajrai et al 2016
Orpheovirus	Orpheovirus IHUMI-LCC2	Fezes de rato	*V. vermiformes*	Ovóide/ 1 µm	1473	Andreani et al 2018
Medusavirus	Medusavírus	Água termal/ Japão	*A. castellanii*	Icosaédrico/ 260 nm	381	Yoshikawa et al 2019

*Taxons reconhecidos pelo ICTV

No Brasil, diversos estudos de prospecção já foram realizados o que culminou no isolamento de centenas de vírus gigantes (ANDRADE *et al.*, 2014; ASSIS *et al.*, 2015; BORATTO *et al.*, 2015; CAMPOS *et al.*, 2014; DORNAS *et al.*, 2015; DORNAS *et al.*, 2016; SILVA *et al.*, 2015). O primeiro vírus gigante isolado no Brasil foi denominado Samba vírus (SMBV), o qual é um mimivírus da linhagem A obtido a partir de amostra de água, coletadas no Rio Negro, na região da Amazônia. Juntamente com o SMBV foi isolado um virófago (um vírus que se multiplica apenas na presença de outro vírus gigante), denominado Rio negro vírus (LA SCOLA et al., 2008; CAMPOS *et al.*, 2014). Posteriormente, outros mimivírus foram isolados a partir de amostras brasileiras, como o Niemeyer vírus, isolado de amostras de água coletadas na lagoa da Pampulha, Belo Horizonte, Minas Gerais (BORATTO *et al.*, 2015); o Kroon vírus, isolado em 2012 de amostra de água de uma lagoa urbana na cidade de Lagoa Santa, Minas Gerais (BORATTO et al., 2018); Amazonia vírus, isolado em 2011 de amostra de água do Rio Negro (ASSIS *et al.*, 2015); e Oyster vírus, isolado a partir de amostras de ostras de uma fazenda situada na costa sul de Florianópolis, Santa Catarina (ANDRADE *et al*, 2014). Outros mimivírus foram isolados de amostras coletadas em diferentes ambientes hospitalares, em Belo Horizonte, Minas Gerais (SILVA *et al*., 2015). Além disso, a circulação de mimivírus foi observada em animais domésticos e silvestres (DORNAS *et al*, 2014). Todos os isolados acima citados foram identificados filogeneticamente como mimivírus do grupo A, assim como o SMBV.

No ano de 2015, Dornas e colaboradores, publicaram um amplo estudo de prospecção de vírus gigantes utilizando amostras de água e solo coletadas na lagoa da Pampulha, Belo Horizonte, Minas Gerais, com o isolamento 69 novos isolados de vírus gigantes, sendo 79,3% destes mimivírus da linhagem A, 4,35% da linhagem C e 1,45% da linhagem B. Além disso, este trabalho relata o isolamento dos primeiros marseillevírus e pandoravírus obtidos a partir de amostras brasileiras (DORNAS *et al*., 2015). No ano seguinte, o segundo marseillevírus brasileiro foi isolado de mexilhões dourados coletados no estado Rio Grande do Sul (DOS SANTOS *et al*., 2016). Em 2018 foram descritos os tupanvírus, estes vírus compõe um novo clado da família *Mimiviridae* e foram isolados a partir de sedimentos de lagoas salinas do Pantanal e solo oceânico coletados a 3000 metros de profundidade na Bacia de Campos no Brasil (ABRAHÃO *et al*., 2018). Os tupanvírus possuem o mais completo arsenal de genes relacionados ao processo

de tradução já descrito. Estes vírus apresentam um capsídeo similar ao observado para os mimivírus, porém possuem uma longa cauda cilíndrica acoplada, resultando em partículas com um tamanho médio de 1.2 µm de extensão, podendo atingir até 2.3 µm devido à plasticidade dessa estrutura (Figura 1F) (ABRAHÃO *et al*., 2018).

Assim como relatado para os isolados brasileiros, outros estudos de prospecção e de metagenômica utilizando amostras de diferentes partes do mundo também obtiveram um maior sucesso de isolamento e/ou detecção de mimivírus, comparado ao de outros grupos de vírus gigantes (KHALIL *et al*., 2016; LA SCOLA *et al.*, 2010; revisado por AHERFI *et al.,* 2016, COLSON *et al*., 2017; MIHARA *et al.,* 2018; SCHULTZ *et al*., 2018). Devido a isso, a família *Mimiviridae* é a família com maior número de isolados entre os vírus gigantes, contendo centenas de representantes. Entretanto, os fatores que favorecem ao maior sucesso de isolamento de mimivírus em relação aos demais vírus gigantes ainda não estão esclarecidos (revisado por AHERFI *et al*., 2016, COLSON *et al*., 2017).

Apesar dos mimivírus serem estudados a mais de uma década, algumas etapas do ciclo de multiplicação destes vírus ainda não foram ou estão parcialmente descritas. Devido ao seu grande tamanho, foi proposto e amplamente aceito que a penetração destes vírus na célula hospedeira ocorre por meio do processo de fagocitose, apesar de nenhum ensaio biológico ter validado esta proposição até o presente o trabalho (MUTSAFI *et al.*, 2010; SUZAN-MONTI *et al.*, 2007). Após a fagocitose, foi demonstrado que o desnudamento se inicia com a abertura do *star-gate*, a membrana interna do vírus se funde com a membrana do fagossomo e a denominada semente viral (genoma e proteínas associadas ao material genético envoltos por uma membrana, formando uma estrutura esférica) é liberada no citoplasma. Os eventos que ocorrem entre a fagocitose e abertura do *star-gate* ainda são pouco conhecidos (MUTSAFI *et al.*, 2010; SUZAN-MONTI *et al.*, 2007). Posteriormente, o DNA viral é liberado no citoplasma e uma fase de eclipse, na qual as partículas virais não são observadas na célula, ocorre (MUTSAFI *et al.*, 2010; SUZAN-MONTI *et al.*, 2007). Durante esta etapa o citoplasma da célula é reorganizado promovendo a formação das fábricas virais (FV), onde o genoma viral é replicado, transcrito e novas partículas são montadas (FRIDMANN-SIRKIS *et al.*, 2016; MUTSAFI *et al.*, 2010).

O envolvimento do núcleo no ciclo de multiplicação destes vírus ainda não está esclarecido. Suzan-Monti e colaboradores em 2007 propuseram que o DNA viral migra para o núcleo celular, onde ocorrem os primeiros ciclos de replicação, depois o genoma volta para o citoplasma, onde os demais ciclos de replicação acontecem. Porém, outros autores defendem que o ciclo de multiplicação destes vírus é exclusivamente citoplasmático, ao demonstrarem que a replicação do genoma é iniciada após a sua liberação da semente viral no citoplasma da célula (MUTSAFI *et al*., 2010).

Kuznetsov e colaboradores em 2013 propuseram um modelo de morfogênese para os mimivírus, baseado em dados de microscopia de força atômica. Neste modelo as etapas de montagem das partículas virais ocorrem de forma sequencial (KUZNETSOV *et al.*, 2013). Primeiramente, a montagem do capsídeo ocorre na periferia das FV e a formação dos capsídeos se inicia pela região do *star-gate*, seguido pelas porções adjacentes. Posteriormente, o DNA viral e outras macromoléculas são inseridos no capsídeo ainda em formação, pelas faces opostas ao *star-gate* (SUZAN-MONTI *et al.*, 2007; KUZNETSOV *et al.*, 2013). Em seguida ocorre a aquisição do genoma, o capsídeo acaba de ser formado e é envolvido por uma camada protaica, na qual as fibrilas, que evolvem externamente toda a partícula viral, se aderem. Além disso, neste modelo, a etapa de aquisição das fibrilas virais ocorre na periferia da célula (KUZNETSOV *et al.*, 2013). Entretanto, a análise detalhada da nossa coleção de imagens de microscopia eletrônica de transmissão e varredura de diferentes etapas do ciclo de multiplicação dos mimivírus, não confirma algumas etapas descritas no modelo acima, como por exemplo, a descrição de que a aquisição de genoma ocorre antes da aquisição fibrilas e que esta última etapa ocorre na periferia celular.

Outro grupo viral, os pandoravírus, também apresenta um ciclo de multiplicação ainda parcialmente caracterizado até o presente trabalho. Philippe e colaboradores em 2013 relataram a descoberta dos pandoravírus e descreveram brevemente algumas etapas do ciclo de multiplicação destes vírus utilizando microscopia eletrônica de transmissão. Eles descreveram que o ciclo se inicia com a com a penetração das partículas virais por fagocitose, e o desnudamento inicia com a fusão da membrana interna viral com a membrana do fagossomo criando um canal na região do ostíolo apical da partícula pelo qual o conteúdo interno viral é

liberado no citoplasma da célula hospedeira. Após o desnudamento, é observada uma fase de eclipse onde as partículas virais não são observadas mais na célula. Durante a etapa de morfogênese foi observado que a partícula começa a ser montada pela região do ostíolo apical e que o capsídeo e o conteúdo interno viral são formados de forma simultânea. Nas etapas mais tardias do ciclo o núcleo celular não é mais visível. O ciclo completo tem duração de 12 horas após infecção (h.p.i) em média (PHILIPPE *et al*., 2013).

Legendre e colaboradores em 2018, relatam o isolamento de outros pandoravírus e também fizeram uma breve análise do ciclo de multiplicação destes novos isolados utilizando microscopia óptica e eletrônica. Eles relataram que o ciclo pode ser mais rápido para alguns isolados como o P. neocaledonia, com duração de 8 h.p.i. Além disso, os autores sugeriram que a liberação viral também pode ocorrer por exocitose por meio de um vídeo (magnificação de 1000X) que mostra partículas sendo liberadas de uma ameba infectada utilizando uma multiplicidade de infecção (MOI) de 1000, 5 h.p.i (LEGENDRE *et al*., 2018). Ainda, sugeriram que as etapas de penetração e desnudamento são semelhantes ao que já foi descrito por Philippe e colaboradores em 2013. Como descrito acima, o ciclo de multiplicação dos mimivírus e pandoravírus ainda estão parcialmente descritos e novos estudos que investiguem mais detalhes acerca das etapas que o compõe são fundamentais na caracterização biológica deste vírus, bem como na compreensão da relação vírus-hospedeiro.

.

2. JUSTIFICATIVA

Nos últimos anos, o estudo dos vírus gigantes tem revelado grandes surpresas quanto à complexidade genômica e estrutural dos vírus. Após o isolamento do APMV, os estudos de prospecção se intensificaram levando a descoberta de novos grupos virais como os marseillevírus, pandoravírus, pithovírus, mollivírus, faustovírus, cedratvírus, tupanvírus, medusavírus. Embora muito se tenha avançado na identificação de novos vírus, pouco se sabe a respeito da diversidade destes vírus e sua distribuição na natureza. Sendo assim o isolamento de novos vírus gigantes é essencial para uma melhor compressão deste universo, possibilitando o avanço da virologia. A descoberta de um novo vírus amplia o conhecimento sobre as interações entre os vírus e seus hospedeiros, sua ecologia, além de revelar possíveis novas características genéticas, fenotípicas e evolutivas. Os mimivírus são os primeiros vírus gigantes de amebas descritos, possuem o maior número de representantes e são os mais bem caracterizados pela literatura. Mesmo assim, no que se refere ao seu ciclo de multiplicação, muitas etapas ainda são pouco compreendidas. A respeito dos pandoravírus também existem poucas descrições em relação ao seu ciclo de multiplicação. Diante do fato de que uma parte essencial para caracterização biológica dos vírus é o entendimento do seu ciclo de multiplicação, esse trabalho foi direcionado não só para a prospecção de novos vírus gigantes, mas também para a aquisição de novas informações acerca da do ciclo de multiplicação dos mimivírus e dos pandoravírus.

3. OBJETIVOS

3.1 Objetivo geral:

Isolamento de novos vírus gigantes a partir de coleções de amostras brasileiras e da Antártica e a caracterização biológica de algumas etapas dos ciclos de multiplicação dos mimivírus e dos pandoravírus.

3.2 Objetivos específicos:

- Isolar novos vírus gigantes a partir de amostras de água, esgoto, solo e fezes coletadas em diferentes regiões brasileiras e da Antártica;
- Identificar os novos isolados molecularmente e/ou morfologicamente;
- Avaliar o efeito dos inibidores farmacológicos citocalasina e bafilomicina na penetração e desnudamento dos mimivírus em células de *A. castellanii*;
- Avaliar a formação do capsídeo dos mimivírus durante a morfogênese;
- Avaliar eventos de aquisição de genoma e fibrilas durante a morfogênese dos mimivírus;
- Avaliar a região em que as fibrilas dos mimivírus são adquiridas durante a morfogênese;
- Avaliar a formação de partículas defectivas durante o ciclo de multiplicação dos mimivírus.
- Caracterizar morfologicamente a fábrica viral de pandoravírus isolados no Brasil;
- Avaliar as modificações nucleares da célula amebiana que ocorrem durante a infecção pelos pandoravírus;
- Avaliar a dinâmica da morfogênese dos isolados de pandoravírus brasileiros;
- Avaliar o efeito do inibidor farmacológico brefeldina na etapa de liberação dos pandoravírus;
- Investigar quais os processos de liberação os pandoravírus utilizam.

4. METODOLOGIA, RESULTADOS E DISCUSSÕES PARCIAIS SERÃO APRESENTADOS NA FORMA DE ARTIGOS PUBLICADOS PRECEDIDOS DE UM BREVE RESUMO.

4.1. Artigo # 1: Ubiquitous Giants: A plethora of giant viruses found in Brazil and Antarctica

Desde a descoberta do APMV, em 2003, os estudos de prospecção viral se intensificaram e outros vírus pertencentes à família *Mimiviridae*, assim como outros grupos virais (marseillevírus, pandoravírus, pithovírus, mollivírus, faustovírus, cedratvírus e o kaumoebavírus) foram descritos. Estes vírus já foram isolados a partir de amostras coletadas em diferentes partes do mundo com condições ambientais distintas. Este trabalho descreve o isolamento de 68 vírus gigantes a partir de uma coleção de 976 amostras coletadas no Brasil e na Antártica. Estes vírus foram isolados após inoculação direta das amostras diluídas em uma cultura de amebas do gênero *Acanthamoeba* contendo antibióticos. A identificação foi feita utilizando coloração hemacolor, PCR e microscopia eletrônica. No total, foram isolados 64 vírus da família *Mimiviridae,* sendo que 15 são da linhagem A, 17 da linhagem B, 2 da linhagem C e 28 de linhagens não identificadas. Estes vírus foram isolados de diferentes tipos de amostras, incluindo água marinha da Antártica, sendo o primeiro relato de isolamento de mimivírus neste ambiente extremo. Além disso, um novo marseillevírus, dois pandoravírus e um cedratvírus foram isolados de amostras de esgoto, adicionando novos membros ao grupo dos NCLDV. Dos diferentes tipos de amostras testadas, a maior diversidade de grupos virais isolados foi obtida de esgoto. Estes resultados reforçam a importância de estudos de prospecção em diferentes amostras ambientais que auxiliam na compreensão da circulação e diversidade destes vírus no ambiente.

Este artigo foi publicado no periódico ***Virology Journal*** em janeiro de 2018.

Após a publicação deste trabalho, as amostras isoladas continuaram a ser analisadas por microscopia eletrônica e alguns isolados foram identificados como outro grupo viral. Os dados de identificação atualizados até o presente trabalho estão disponíveis na tabela suplementar do anexada após o artigo.

Andrade et al. Virology Journal (2018) 15:22
DOI 10.1186/s12985-018-0930-x

Virology Journal

RESEARCH Open Access

Ubiquitous giants: a plethora of giant viruses found in Brazil and Antarctica

Ana Cláudia dos S. P. Andrade[1], Thalita S. Arantes[1], Rodrigo A. L. Rodrigues[1], Talita B. Machado[1], Fábio P. Dornas[1], Melissa F. Landell[2], Cinthia Furst[3], Luiz G. A. Borges[4,5], Lara A. L. Dutra[6], Gabriel Almeida[6], Giliane de S. Trindade[1], Ivan Bergier[7], Walter Abrahão[8], Iara A. Borges[1], Juliana R. Cortines[9], Danilo B. de Oliveira[10], Erna G. Kroon[1] and Jônatas S. Abrahão[1*]

Abstract

Background: Since the discovery of giant viruses infecting amoebae in 2003, many dogmas of virology have been revised and the search for these viruses has been intensified. Over the last few years, several new groups of these viruses have been discovered in various types of samples and environments.In this work, we describe the isolation of 68 giant viruses of amoeba obtained from environmental samples from Brazil and Antarctica.

Methods: Isolated viruses were identified by hemacolor staining, PCR assays and electron microscopy (scanning and/or transmission).

Results: A total of 64 viruses belonging to the *Mimiviridae* family were isolated (26 from lineage A, 13 from lineage B, 2 from lineage C and 23 from unidentified lineages) from different types of samples, including marine water from Antarctica, thus being the first mimiviruses isolated in this extreme environment to date. Furthermore, a marseillevirus was isolated from sewage samples along with two pandoraviruses and a cedratvirus (the third to be isolated in the world so far).

Conclusions: Considering the different type of samples, we found a higher number of viral groups in sewage samples. Our results reinforce the importance of prospective studies in different environmental samples, therefore improving our comprehension about the circulation anddiversity of these viruses in nature.

Keywords: Giant viruses, Prospection, Brazil, Antarctica, Pandoravirus, Cedratvirus, Marseillevirus, Mimivirus

Background

The discovery of *Acanthamoeba polyphaga mimivirus* (APMV) in 2003, the first isolated giant virus infecting amoebas, interested the scientific community due to its size and genome content, which culminated in the search for and isolation of new giant viruses [1, 2]. The giant amoebal viruses have many phenotypic and genomic features which had never been seen in other viruses before, like large viral particles presenting up to 1.5 μm in length and large double-stranded DNA genomes ranging from 350 kb in *Marseilleviridae* members to 2500 kb for pandoravirus [3, 4]. These genes encode many hypothetical proteins, uncharacterized, or with functions that have never orrarely been observed before in other viruses, such as those related to translation and DNA repair [5–7]. Common characteristics shared by giant and large DNA viruses permitted their incorporation into a supposedly viral monophyletic group, named nucleocytoplasmic large DNA viruses (NCLDV), created in 2001 [8]. When the NCLDV group was proposed, it was composed of families *Poxviridae (*e.g. *Vaccinia virus, Crocodilepox virus), Asfarviridae (*e.g. *African swine fever virus*) *Iridoviridae* (e.g. *Frog virus 3)* and *Phycodnaviridae* (e.g. *Emiliania huxleyi virus 86, Aureococcus anophagefferens virus)* [8].

Subsequently, viruses belonging to the Mimiviridae, Marseilleviridae, Ascoviridae family and also the pandoravirus, faustovirus, pithovirus, mollivirus, kaumoebavirus, cedratvirus and pacmanvirus were also incorporated to NCLDV group [9–17]. Recent prospective studies have

* Correspondence: jonatas.abrahao@gmail.com
[1]Laboratorio de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
Full list of author information is available at the end of the article

shown that giant viruses are ubiquitous, as are their protozoa hosts [2, 18, 19]. The use of high-throughput techniques and different species of amoebae in culture for viral isolation has allowed the discovery of a large variety of new viruses and new lineages in recent years. They have been detected and/or isolated in all continents of Earth. Metagenomic studies have indicated an outstanding profile of giant virus distribution and diversity in natural environments and organisms, including water, soil, invertebrates and mammals [14, 20–27]. It is important to note that mimiviruses and marseilleviruses have also been isolated from human samples, raising questions abouttheir possible role as pathogenic agents of diseases, but this possibility still under investigation, and these viruses may be components of healthy humans virome [24, 25, 28–31].

Despite the advances made in the techniques used to isolate new giant viruses, which have increased the success of detection and the isolation of these viruses in different environments around the world, the diversity, distribution and role of these viruses in nature is still far from completely understood. Therefore, in order to better understand the diversity and distribution of giant viruses in the environment, this work aimed at the isolation and identification of giant viruses obtained from clinical and environmental samples from different regions of Brazil and Antarctica. A total of 976 samples were analyzed and 68 viruses were isolated. Taken together, our results reinforcethat giant viruses, in particular mimiviruses, are ubiquitous and may play an important role in the control of amoebal populations, both in natural and anthropogenic-affected environments.

Methods

Samples collection and treatment

In this work, a collection composed of 976 clinical and environmental samples was analyzed: 495 soil samples (mean weight was 3 g of each sample), 124 water samples (10 mL of each sample), 140 sewage samples (10 mL of each sample), 200 human nasopharyngeal aspirate samples (1,5 mL of each sample) and 17 capybara samples (mean weight was 2 g of each sample) (Table 1 and Fig. 1).

Table 1 Collections and locations of samples analyzed

Collections	Type of sample	Collection site	Date of collection
Serra do Cipó			
13 samples	Freshwater	Serra do Cipó, MG, Brazil	Jan.2015
47 samples	Soil	Serra do Cipó, MG, Brazil	Jan. 2015
Sewage creeks Pampulha			
110 samples	Sewage	Pampulha Creeks, Belo Horizonte, MG, Brazil	Oct.2016
Farm Sewage			
30 samples	Sewage	Itaúna, MG, Brazil	Nov. 2016
Water treatment station			
50 samples	Freshwater	COPASA, Belo Horizonte, MG, Brazil	Dec. 2016
Antarctic			
7 samples	Marine Water	Antarctic	Dec. 2014
Capybara Stool			
17 samples	Stool	Serra do Cipó, MG, Brazil Pampulha, MG, Brazil Serro, MG, Brazil Pantanal, MS, Brazil	Dec. 2012 Dec. 2012 Dec. 2012
Minas Gerais Soil			
470 samples	Soil	MG, Brazil	Jan. 2014
Pantanal soil			
12 samples	Soil	Pantanal, MT, Brazil	Mar. 2015
Human nasopharyngeal aspirate			
200 samples	Human nasopharyngeal aspirate	Laboratório Central do Estado do Rio Grande do Sul, Brazil	Nov. 2014
Bromeliads Water			
10 samples	Freshwater	Maceió, AL, Brazil	Set. 2015
Mangrove water			
10 samples	Mangroove water Marine water	ES, Brazil	Feb. 2015

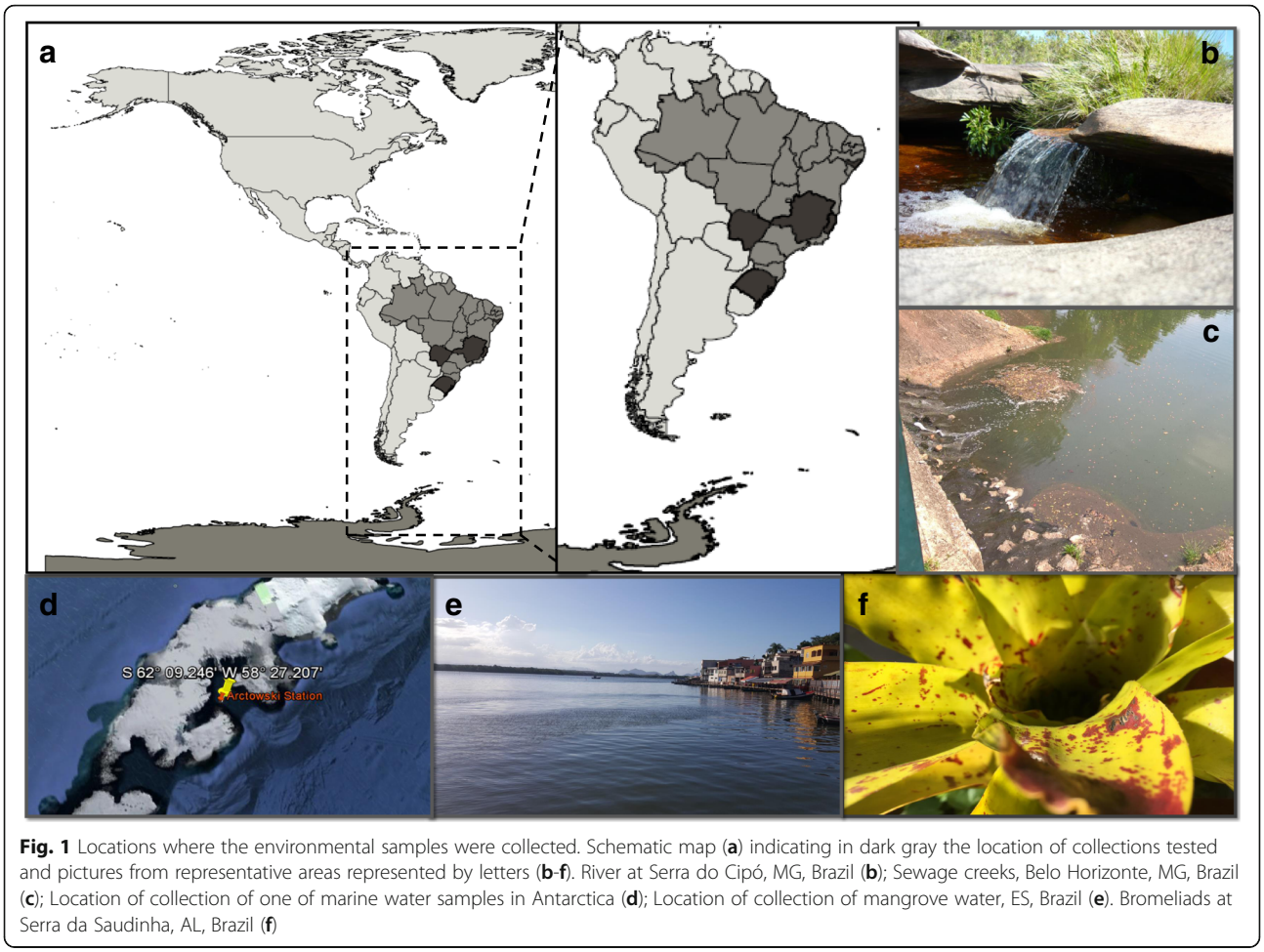


Fig. 1 Locations where the environmental samples were collected. Schematic map (**a**) indicating in dark gray the location of collections tested and pictures from representative areas represented by letters (**b**-**f**). River at Serra do Cipó, MG, Brazil (**b**); Sewage creeks, Belo Horizonte, MG, Brazil (**c**); Location of collection of one of marine water samples in Antarctica (**d**); Location of collection of mangrove water, ES, Brazil (**e**). Bromeliads at Serra da Saudinha, AL, Brazil (**f**)

All collections were collected in different locations using sterile tubes.

The samples of human nasopharyngeal aspirate were used under approval of the ethics committee of Universidade Federal de Ciências da Saúde de Porto Alegre (protocol number 1774/12, register 928/12). After collection, all samples were stored at 4 °C until inoculation procedures were performed.

Initially, the samples were divided into two groups, one with sediment-free water, including human clinical samples and other with a high concentration of sediment and soil. Samples with only water and no sediment were directly inoculated onto amoebalcultures. The soil samples were transferred to conical tubes of 15 mL and treated with 5 mL of phosphate buffered saline (PBS). The system was left for 24 h for sediment decantation and then the supernatants were collected and inoculated onto amoebal cultures.

Culture procedures

For viral isolation, we used *Acanthamoeba polyphaga* (ATCC 30461), *Acanthamoeba castellanii* (ATCC 30234) kindly provided by the Laboratório de Amebíases (Departamento de Parasitologia, ICB/UFMG) and *Vermamoeba vermiformis* (ATCC CDC19), kindly provided by Professor Bernard La Scola from Aix Marseille University. Amoeba were grown in 75 cm^2 Nunc™ Cell Culture Treated Flasks with Filter Caps (Thermo Fisher Scientific, USA) with 30 mL of peptone-yeast extract-glucose (PYG) medium supplemented with 0,14 mg/mL penicillin (Sigma-Aldrich, USA), 50 mg/mL gentamycin (Thermo Fisher Scientific, USA) and 2.5 mg/mL amphotericin (Bristol-Myers- Squibb, New York, USA) at 32 °C.

For co-culture, amoeba were re-suspended in 10 mL of PYG supplemented with an antibiotic mix containing 0,004 mg/mL ciprofloxacin (Cellofarm, Brazil), 0,004 mg/mL vancomycin (Sigma-Aldrich, U.S.A), and 0,020 mg/mL doxycycline (Sigma-Aldrich, U.S.A). The suspension was then diluted 1:10 in PBS and then inoculated in96-well plates containing 4×10^4 cells per well. The plates were incubated for 7 days at 32 °C and observation of the cytopathic effect was done daily using an inverted optical microscope. The well contents were then collected, frozen and thawed three times to lyse the bacterial and fungal

cells that may be present in the samples and thereby decrease the chance of co-culture contamination and also helps release the viruses of amoeba cells not lysed. Posteriorly, the samples were re-inoculated for two new sub-cultures on fresh amoeba, as described above (blind passages). The contents of wells with cytopathic effect were collected and inoculated in a new 25 cm^2 Nunc™ Cell Culture Treated Flasks with Filter Caps (Thermo Fisher Scientific, USA) culture containing 1 million cells, the cytopathic effect was confirmed and this culture was centrifuged 10,000 rpm for 10 min (Centrifuge Sigma 1–14) for lysate clearance and were further analyzed for giant viruses. Negative controls with no sample inoculated amoeba were used in all microplates.

DNA extraction and PCR

After the identification of cultures with a cytopathic effect, screening was done to identify which giant virus was present in samples using PCR with specific targets for some giant virus groups (Table 2). For this, 200 μL of each positive suspension was used for DNA extraction. DNA was extracted using the phenol-chloroform method [32] and used at the concentration 50 μg/μg as a template for PCR assays. The genes targeted in the PCR assays were: helicase of mimivirus lineage A; DNA polymerase B of mimivirus lineage B; DNA polymerase B of mimivirus lineage C; the major protein of the capsid of the family *Mimiviridae* (generic reaction targeting lineages A, B and C), *Marseilleviridae*, pandoravirus and cedratvirus. The primers were designed using a freely available primer design tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) at the National Center for Biotechnology Information, U.S.A (NCBI); the sequences are described in Table 2. The primers and reactions were designed and standardized considering all analyzed viruses available on GenBank to avoid cross-amplification. PCR assays were performed using 1 μL of extracted DNA (~ 50 nanograms) in an amplification reaction mix containing 5 μL of SYBR Green Master Mix and 0.4 μL (10 μM) of forward and reverse primers. The final volume of the reaction was adjusted with ultrapure waterto 10 μL. The conditions of the StepOne thermal cycler reactions (Applied Biosystem, USA) were: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min, which was followed by a final step of 95 °C for 15 s, 60 °C for 1 min and 95 °C for 15 s. Positive samples in the PCR were those that amplified, showing the specific melting temperature, using the primers listed in Table 2, whereas the negative samples did not amplify in the PCR. As negative controls we used DNA extracted from non-inoculated amoebas with purified viruses or samples, and as a positive control we used DNA from amoebae infected with purified virus. Samples that were not possible to identify using the PCR assay were identified by electron microscopy and/or hemacolor staining.

Sequencing validation and phylogeny

Four isolates were selected for sequencing validation. The genome of two pandoraviruses, the cedratvirus and one mimiviruses of lineage B positive samples were sequenced using the Illumina MiSeq instrument (Illumina Inc., San Diego, CA, USA) with the paired-end application. The sequenced reads were imported to CLC_Bio software and assembled into contigs by the de novo method. The prediction of open reading frame (ORF) sequences was carried out using the Fgenes V tool. ORFs smaller than 100aa were excluded from the annotation. Paralogous groups of genes were predicted by OrthoMCL program. The ORFs were functionally annotated using similarity analyses with sequences in the NCBI data base using BLAST tools. One fragment of 327 amino acid of DNA polymerase B gene sequence of the samples was aligned with sequences from other giant viruses, previously deposited in GenBank, using the ClustalW program. After the alignment analysis, phylogeny reconstruction was performed using the Neighbor-joining method implemented by the MEGA7 software.

Viral stock production and titration

For seed pool production, *A. castellanii* or *A. polyphaga* cells were cultivated and infected with 500 μL of isolates. After observation of a cytopathic effect, the titer was

Table 2 Primer sequences used for specific PCR

Target genes	Forward sequence	Reverse sequence
Helicase of mimivirus lineage A	5′-ACCTGATCCACATCCCATAACTAAA-3′	5′-GGCCTCATCAACAAATGGTTTCT-3′
DNA polymerase beta of mimivirus lineage B	5′-AGTTCATCCGCACTTGGAGA-3′	5′-TCAACGGATAAAATCCCTGGTACT-3′
DNA polymerase beta of mimivirus lineage C	5′- TCCGAATTCTATGAGGGAGAGA-3′	5′-TGTTCCTTTTTGGGAGAACCA-3′
Main protein of the capsid of the family *Mimiviridae*	5′-ACTTTATTATCATTATCAGCGAATA-3′	5′-GCTCTTAACCCTGAAGAACA-3′
Main protein of the capsid of the family *Marseilleviridae*	5′-CTTTTGCACCTGCTTCATGA-3′	5′-GCGGTAACCCTCCCACTTAT-3′
Main protein of the capsid of pandoravirus	5′-GGATGGCTCGACGTCTCTT-3′	5′-CCTYGGTRAGCAMAGGCAAC-3′
Main protein of the capsid of cedratvirus	5′- AGAGTATGCTCGCAACCACC-3′	5′-CACGTTAAGGCCGGGGTAAT −3′

obtained by end-point method [33]. Stocks were kept at − 80 °C freezers.

Hemacolor staining

A. castellanii and *A. polyphaga* cells were infected with isolates at a M.O.I of 0.01 following the procedures described above. After approximately 18 h, amoeba became rounded, so 10 μL of the previously inoculated suspension was spread on a histological slide and fixed with methanol. The virus factories and viral particles were observed after hemacolor (Renylab, Brazil) or crystal violet (Labsynth, Brazil) staining, respectively. After, slides were analyzed under an optical microscope (OlympusBX41, Japan) with 1000X zoom.

Electron microscopy

For transmission electron microscopy (TEM), *A. castellanii* and *A. polyphaga* cells were cultivated until the observation of 80–90% confluence and infected with the isolates in an M.O.I of 0.01. The samples were prepared as described previously [34]. Briefly, 12 h post-infection, when approximately 50% of the cells were presenting a cytopathic effect, the medium was discarded and the monolayer gently washed twice with 0.1 M phosphate buffer. Glutaraldehyde 2.5% (*v*/v) was added to the system, followed by incubation for 1 h at room temperature for fixation. The cells were then collected, centrifuged at 3000 rpm for 10 min, the medium discarded and the cells stored at 4 °C in phosphate buffer until electron microscopy analyses.

For the scanning electron microscopy (SEM) assay, the isolates were prepared onto round glass blades covered by poly-L-lysine and fixed with glutaraldehyde 2.5% in 0.1 M cacodylate buffer for 1 h at room temperature. Samples were then washed three times with 0.1 M cacodylate buffer and post-fixed with 1.0% osmium tetroxide for 1 h at room temperature. After a second fixation, the samples were washed three times with 0.1 M cacodylate buffer and immersed in 0.1% tannic acid for 20 min. Samples were then washed in cacodylate buffer and dehydrated by serial passages in ethanol solutions with concentrations ranging from 35% to 100%. They were dried at the critical CO_2 point, transferred into stubs and metalized with a 5 nm gold layer. The analyses were completed with scanning electronic microscopy (FEG Quanta 200 FEI) at the Center of Microscopy of UFMG, Brazil.

Results

Here, we report the screening of 976 environmental and clinical samples collected between 2014 and 2017 and the isolation of 68 giant viruses (6.97% isolation rate). Among all of the isolated viruses, 17 (25%) were isolated in *A. polyphaga* and 51 were isolated in *A. castellanii* (75%). No virus was isolated in *V. vermiformes* (Additional file 1: Table S1).

The PCR, hemacolor staining and electron microscopy assays showed that 22 samples were positive for mimivirus lineage A, 17 were positive for mimivirus lineage B and 2 were positive for mimivirus lineage C. In addition, 2 samples were positive for pandoravirus, 1 for cedratvirus and 1 for marseillevirus (Fig. 2). Twenty-three other samples were identified as mimiviruses by PCR (capsid gene, generic reaction), by hemacolor staining or by electron microscopy, but it was not possible to discriminate the lineage of these viruses using the specific PCR (Additional file 1: Table S1).

Twelve samples were positive in PCR for mimivirus lineage B and for marseillevirus. Four isolates were selected for genome sequencing and phylogenetic analyzes performed with the DNA polymerase gene of these viruses confirmed the identification by PCR (Fig. 3). In order to investigate the occurrence of co-infections, these samples were analyzed via hemacolor staining, and two of them were randomly selected for diagnosis via TEM. The samples tested showed only particles with morphology similar to the mimivirus, showing no marseillevirus-like particles (Additional file 1: Table S1). In addition, no marseillevirus-like factories were observed by hemacolor staining, just mimiviruses-like ones.

The highest isolation percentages (27.42%) were obtained from the water samples, with 34 isolates from 124 samples. Of these, 4 were isolated from 7 seawater samples (57.14% isolation rate) and 30 were isolated from 117 freshwater samples (isolation rate of 25.64%) (Fig. 2).

In addition, with an isolation success of 18.57%, 26 viruses were obtained from 140 sewage samples, followed by samples of capybara feces (5.88%), with an isolate obtained from 17 samples, and soil samples (1.41%), with 7 isolates from 495 samples. In addition, 200 samples of human nasopharyngeal aspirate were tested and no isolates were obtained from these samples (Fig. 2).

Although water samples have shown the highest number of isolated virus, sewage samples presented the highest diversity of viruses groups isolated (Fig. 2). In the fresh and marine water samples, only *Mimiviridae* family viruses (12 of lineage A, 13 of lineage B and 9 unidentified) were identified, while besides *Mimiviridae* (9 of lineage A, 2 of lineage C and 11 unidentified), 1 marseillevirus, 1 cedratvirus and 2 pandoraviruses (Fig. 2) were found in the sewage samples. Soil and stool samples also showed only viruses of *Mimiviridae* family (4 of lineage B and 3 unidentified in soil, and 1 of lineage A in stool) (Fig. 2).

Comparing the percentage of isolates per collection region, we can observe that the isolation was higher in the Antarctica collection with a 57% isolation success rate; however, this collection has few samples (4 isolates

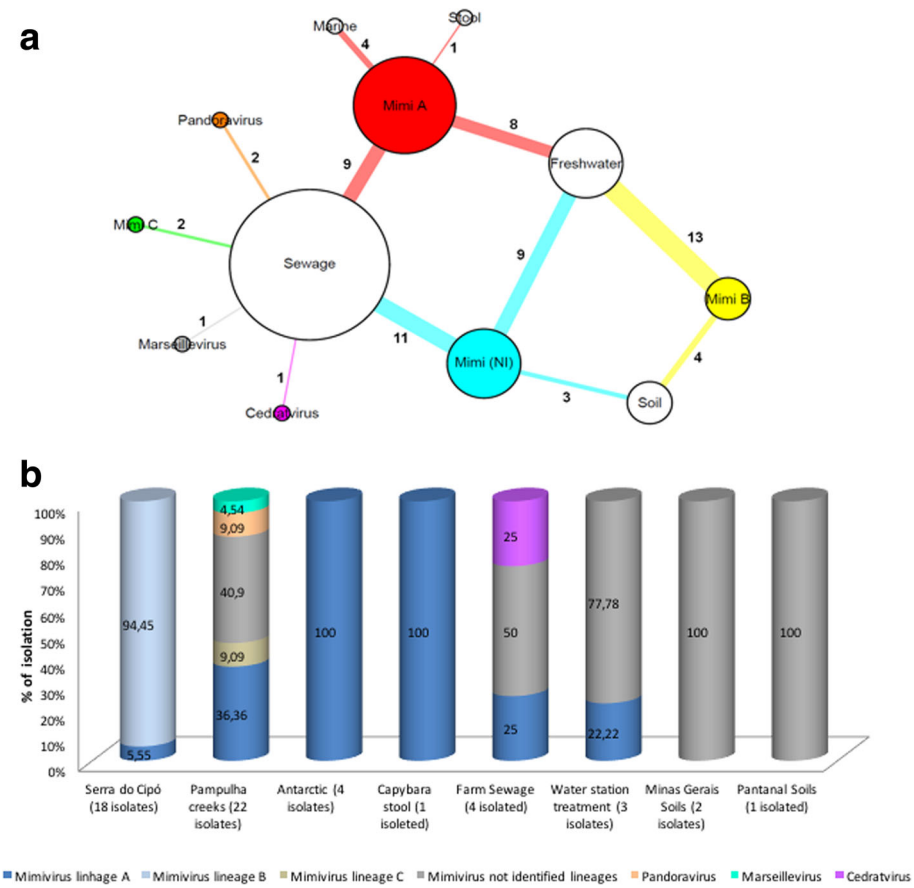


Fig. 2 Diversity of isolated giant virus by type of sample and collections. Network graph showing the viral groups isolated and identified by PCR and electron microscopy assays in different samples. Each node represents a type of sample (white nodes) or viral group (colored nodes). The node diameter is proportional to the edge degree. The numbers of isolated viruses in each sample are shown on the respective edge. The layout was generated using a force based algorithm followed by manual rearrangement for a better visualization of the connections (**a**). A total of 7 viral groups are represented. Isolation rate of each virus groups by collections (**b**)

from 7 samples). The collections of Serra do Cipó and sewage creeks appear with 30% (18 isolates from 60 samples) and 20% (22 isolates from 110 samples) positivity, respectively. The collections of farm sewage and Pantanal soils presented a percentage of isolations of 13.33% (4 isolates from 30 samples) and 8.33% (1 isolate from 12 samples), respectively (Fig. 2).

The collections of water station treatment, capybara stool and Minas Gerais soils, showed percentages of isolation of 8% (4 isolates from 50 samples), 5.88% (1 isolate from 17 samples) and 0,4% (2 isolates from 470 samples), respectively. Collections from bromeliad, mangrove, and human nasopharyngeal aspirate, showed no viral isolates (Fig. 2). The collection of creek sewages showed the greatest viral diversity, with isolates of *Mimiviridae*, *Marseilleviridae* and *Pandoravirus* groups; followed by the collection of farm sewage, with isolates of mimivirus and cedratvirus. In the remaining collection, only mimiviruses were identified (Fig. 2).

Electron microscopy assays showed that two isolated samples of Mergulhão (Fig. 4b-c) and Bom Jesus (Fig. 4e) sewage creek show pandoravirus-like morphology, with particles having an average length of 1 μm, as described by Philippe and colleagues in 2013. Antarctica isolates showed a mimivirus-like morphology, with particles of about 750 nm, as described by La Scola and colleagues in 2003 (Fig. 4d and g). SEM analyses of a sample of sewage farm collection showed cedratvirus-like morphology with particles of approximately 1.2 μm, as described by Andreani and colleagues in 2016 (Fig. 4a and d). The images obtained from another isolate from Bom Jesus creek showed particles with marseillevirus-like morphology, apparently with icosahedral symmetry and dimensions of about 200 nm (Fig. 4f).

Discussion

The search for giant viruses in environmental and clinical samples from different regions of Brazil and

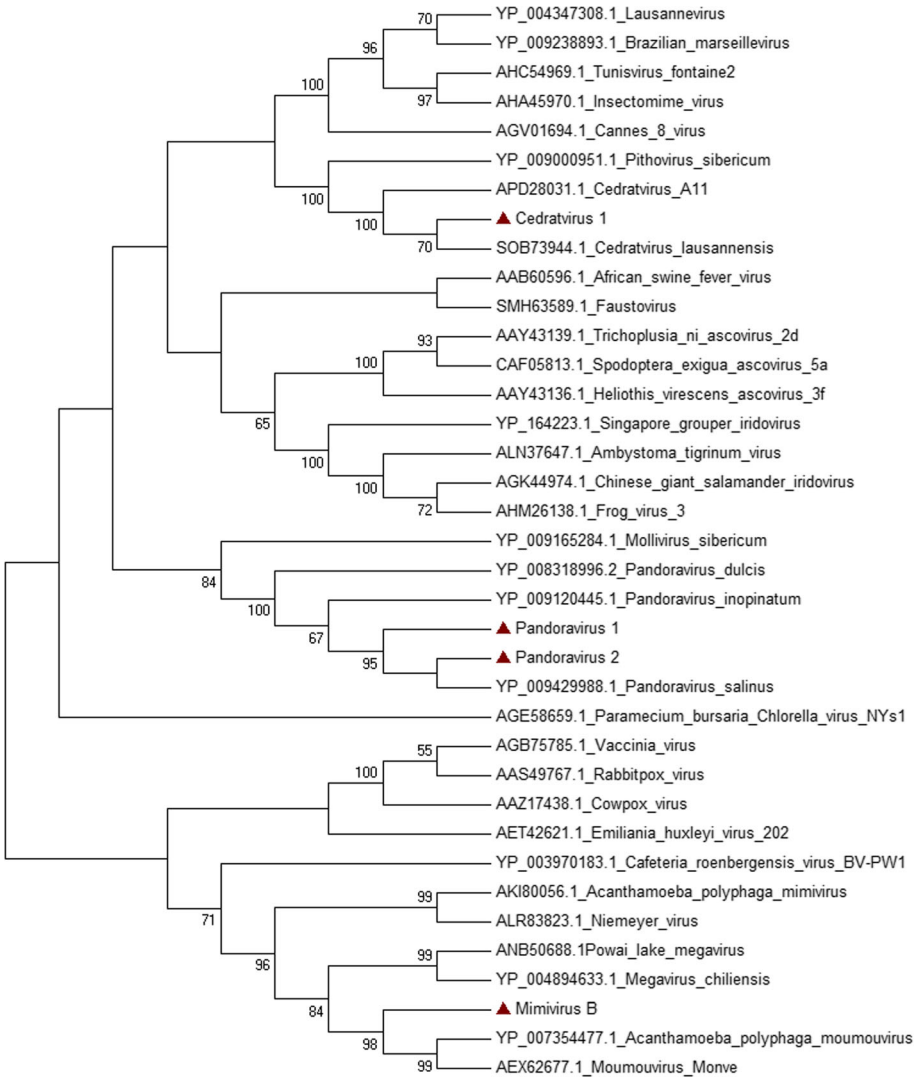


Fig. 3 Phylogenetic tree of isolates. A neighbor-joining phylogenetic tree constructed using a 327 amino acid fragment of the DNA polymerase B gene. Tree was constructed by using MEGA version 7.0 (www.megasoftware.net) on the basis of the amino acids sequences with 1000 bootstrap replicates. Bootstrap values > 40% are shown. Nucleotide sequences were obtained from GenBank. The isolates are highlighted with red triangle. Scale bar indicates rate of evolution

Antarctica resulted in 68 isolates, reinforcing the results obtained in other prospective studies involving environmental Brazilian samples, in which a large variety of giant viruses, specially mimiviruses, were isolated [18, 35–39]. The present work corroborates those studies, since 64 out of 68 viruses isolated (94.11%) were identified as mimiviruses; however, for the first time many mimiviruses of lineages B and C were isolated in Brazil.

Although Brazil is one of the most exploited countries regarding the presence of giant viruses, only two viruses of the *Marseilleviridae* family had been isolated in this territory to date. Brazilian marseillevirus and Golden marseillevirus presented a high genomic diversity, thus suggesting that these isolates form two new lineages within the family [37, 38]. This study presents the third marseillevirus isolated from Brazil. The genomic characterization of this isolate (in progress) may expand even more the plethora of marseillevirus lineages.

Regarding pandoraviruses, since their discovery in 2013, only 4 isolates have been described worldwide [18, 40–42]. These viruses form a new group among the NLCDVs, known as new TRUC (an acronym for Things Resisting Uncompleted Classification) members [10]. Here, we add two members to this club, providing the possibility of a wider study of this virus biology. In addition, this study reports for the first time the isolation of a giant virus from capybara feces. This type of sample had not yet been explored for the presence of giant viruses although DNA from

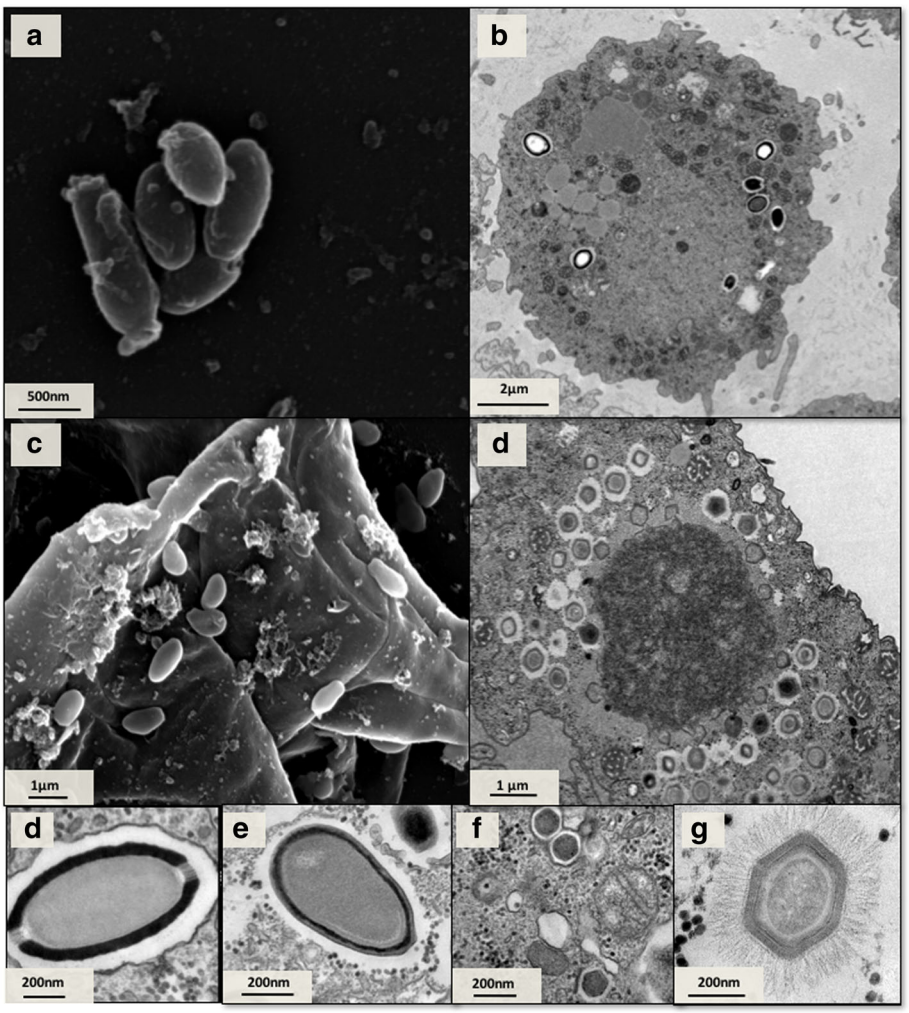


Fig. 4 Electron microscopy images of viruses isolated. SEM of *Cedratvirus* isolated from sewage farm of MG (**a**) TEM (**b**) and SEM (**c**) of *Pandoravirus* isolate from Mergulhão sewage creek. TEM of *Mimivirus* isolated from Antarctica (**c**) TEM of *Cedratvirus* isolated from sewage farm of MG (**d**) TEM of *Pandoravirus* isolate from Bom Jesus sewage creek (**e**) marseillevirus isolated from Bom Jesus sewage creek (**f**). TEM mimivirus particle detail that was isolated from Antarctica (**g**). Scale Bars: (**a**-**d**) 500 nm; (**e**) 50 nm

poxvirus, another member of the NCLDV group, was previously detected in this collection [43].

The isolation and detection rates of giant viruses vary in the different studied samples, with noisolation in human nasopharyngeal aspirate samples and higher rates in water and sewage, followed by stool and soil samples. These results corroborate other studies in which giant viruses are more abundant in water and sewage than in soil samples [18, 19, 34, 39]. Furthermore, we demonstrated that giant viruses are not commonly found in nasopharyngeal clinical samples, as reported elsewhere [44–46]. Considering the amoebas used in this study, *A. castellanii* was shown to be more effective in the isolation of a greater diversity of giant viruses, as demonstrated by Dornas and colleagues in 2015.

The difficulty in amplifying some preserved lineage-specific regions by PCR can be explained by the high genetic diversity among these viruses [36]. This may be one of the reasons why we could not identify the strains of all mimiviruses isolated in this work. It is also important to consider that among these, there may be new strains which have not yet been described. In addition, two samples that were PCR positive for mimivirus lineage B and marseillevirus revealed only mimiviruses particles or factories when analyzed by electron microscopy or hemacolor staining. This finding also reinforces the importance ofusing a set of techniques for the identification of giant viruses, as performedin this study.

Metagenomic studies have indicated that the presence of the giant virus gene marker is common in all continents including Antarctica, a region with extreme environmental conditions [20, 21, 47, 48]. Virophages have already been isolated from this region, which is an additional indicative of the presence of giant viruses

[22, 49]. However, to our knowledge, there has been no description of mimiviruses isolated in this continent to date. Nevertheless, we report the first giant amoebal viruses in Antarctica, confirming some previous expectations and the ubiquity of these microorganisms.

Altogether, our results lead us one step further into knowledge about the giant virus diversity and ecology, but important questions were raised. What could be the role of giant viruses in an extreme environment such as Antarctica? Will the host spectrum of these viruses be the same, or are they capable of infecting other more well-adapted hosts at extreme conditions? In-depth investigations regarding genetic and biological aspects of these isolates might provide some answers. Moreover, new prospecting studies, exploring different isolation strategies in environments that have never been explored around the globe, will bring insights about the ecology of giant viruses and completely new NCLDV members could be brought to light, boosting our knowledge about the diversity of this complex group within the virosphere.

Conclusions

This work presented the isolation of different giant virus species from the prospecting study of a large collection of environmental samples, providing the isolation of viruses never previously isolated in Brazil and Antarctica. The findings of this study reinforce the idea that giant viruses are ubiquitous and open the door to further study of the biology of these isolates, which contributes to an understanding of the diversity of these viruses.

Additional file

Additional file 1: Table S1. Identification and locations of viruses isolated. (DOCX 34 kb)

Abbreviations

AL: Alagoas; APMV: *Acanthamoeba polyphaga mimivirus*; ES: Espírito Santo; LACEN/RS: Laboratório Central do Estado do Rio Grande do Sul; M.O.I: Multiplicity of infection; MG: Minas Gerais; NCBI: National Center for Biotechnology Information; NCLDV: Nucleocytoplasmic large DNA viruses; PBS: Phosphate buffered saline; PCR: Polymerase chain reaction; SEM: Scanning electron microscopy; TEM: Transmission electron microscopy

Acknowledgments

We thank our colleagues from Gepvig and the Laboratório de Vírus fortheir excellent technical support. We thank CNPq, CAPES and FAPMEIG for scholarship and the Center of Microscopy of UFMG. To Drs. Ulysses Lins, Fernanda de Ávila Abreu, Pedro Leão, Tatiana Schaffer Gregianini, Ana Beatriz Gorini da Veiga, for crucial participation on sample collection. To the financial and logistic support from the Brazilian Antarctic Program, PROANTAR, as part of the IPY Activity no. 403 'MIDIAPI Microbial Diversity of Terrestrial and Maritime ecosystems in Antarctic Peninsula' (520194/2006-3) is acknowledged. The Brazilian National Science and Technology Instituteon Antarctic Environmental Research (INCT-APA, CNPq 574018/2008) is also acknowledged. JSA, EGK, GST are CNPq researchers.

Funding

The scholarships were provided by CNPq, CAPES and FAPMEIG. The financial and logistic support were provided by The Brazilian National Science and Technology Institute on Antarctic Environmental Research (INCT-APA, CNPq 574018/2008) and Brazilian Antarctic Program, PROANTAR, as part of the IPY Activity no. 403 'MIDIAPI Microbial Diversity of Terrestrial and Maritime ecosystems in Antarctic Peninsula' (520194/2006–3). JSA, EGK, GST are CNPq researchers.

Availability of data and materials

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Authors' contributions

ACSPA: performed experiments, wrote the manuscript, microscopical assays, designed the experiments; TSA: performed experiments; RALR: performed experiments, TBM: performed experiments; IAB: performed experiments; LALD: performed experiments FPD: microscopical assays; JSA: designed the experiments; EGK: designed the experiments; DBO: designed the experiments; GST: samples collections; GA: samples collections; JRC: samples collections; WA: samples collections; IB: samples collections; LGAB: samples collections; CF: samples collections; MFL: samples collections. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Two hundred samples of human nasopharyngeal aspiratewere also kindly provided by Laboratório Central do Estado do Rio Grande do Sul (LACEN/RS) (Table 1). These samples were used underapproval of the ethics committee of Universidade Federal de Ciências da Saúde de Porto Alegre (protocol number 1774/12, register 928/12).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Laboratorio de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. [2]Laboratório de Diversidade Molecular, Instituto de Ciências Biológicas e da Saúde, Universidade Federal de Alagoas, Maceió, Brazil. [3]Departamento de Patologia, Universidade Federal do Espírito Santo, Maruípe, Brazil. [4]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [5]Instituto do Petróleo e dos Recursos Naturais (IPR), Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil. [6]Department of Biological and Environmental Sciences, University of Jyvaskyla, Jyvaskyla, Finland. [7]Embrapa Pantanal, Corumbá, Brazil. [8]Universidade Federal de Viçosa, Viçosa, Brazil. [9]Departamento de Virologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [10]Faculdade de Medicina, Universidade Federal do dos Vales do Jequitinhonha e Mucuri, Diamantina, Brazil.

Received: 17 August 2017 Accepted: 12 January 2018
Published online: 24 January 2018

References

1. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, et al. A giant virus in amoebae. Science. 2003;299:2033.
2. Aherfi S, Colson P, La Scola B, Raoult D. Giant viruses of amoebas: an update. Front Microbiol. 2016;7:1–14.
3. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, et al. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci. 2009;106:21848–53.
4. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci. 2014;111:4274–9.

5. Abrahão JS, Araújo R, Colson P, La Scola B. The analysis of translation-related gene set boosts debates around origin and evolution of mimiviruses. PLoS Genet. 2017;16:1–12.
6. Colson P, La Scola B, Levasseur A, Caetano-Anollés G, Raoult D. Mimivirus: leading the way in the discovery of giant viruses of amoebae. Nat Rev Microbiol. 2017;4:243–54.
7. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of mimivirus. Science. 2004;306:1344–50.
8. Iyer LM, Aravind L, Koonin EV. Common origin of four diverse families of large eukaryotic DNA viruses. J Virol. 2001;75:11720–34.
9. Sharma V, Colson P, Pontarotti P, Raoult D. Mimivirus inaugurated in the 21st century the beginning of a reclassification of viruses. Curr Opin Microbiol. 2016;31:16–24.
10. Sharma V, Colson P, Chabrol O, Scheid P, Pontarotti P, Raoult D. Welcome to pandoraviruses at the "Fourth TRUC" club. Front Microbiol. 2015;6:1–11.
11. Sharma V, Colson P, Chabrol O, Pontarotti P, Raoult D. Pithovirus sibericum, a new bona fide member of the "Fourth TRUC" club. Front Microbiol. 2015;6:1–9.
12. Benamar S, DGI R, Bandaly V, Labas N, Raoult D, La Scola B. Faustoviruses: comparative genomics of new megavirales family members. Front Microbiol. 2016;7:1–9.
13. Andreani J, Bou Khalil JY, Sevvana M, Benamar S, Di Pinto F, Bitam I, et al. Pacmanvirus, a new giant icosahedral virus at the crossroads between *Asfarviridae* and Faustoviruses. J Virol. 2017;91:212–7.
14. Legendre M, Lartigue A, Bertaux L, Jeudy S, Bartoli J, Lescot M, et al. In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting *Acanthamoeba*. Proc Natl Acad Sci. 2015;112:5327–35.
15. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, Raoult D, et al. Kaumoebavirus, a new virus that clusters with Faustoviruses and Asfarviridae. Viruses. 2016;8:1–10.
16. Andreani J, Aherfi S, JYB K, Di Pinto F, Bitam I, Raoult D, et al. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. Viruses. 2016;8:1–11.
17. Bertelli C, Mueller L, Thomas V, Pillonel T, Jacquier N, Greub G. Cedratvirus lausannensis - digging into Pithoviridae diversity. Environ Microbiol. 2017;10:4022–34.
18. Dornas FP, JYB K, 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;6:1–9.
19. La Scola B, Campocasso A, N'Dong R, Fournous G, Barrassi L, Flaudrops C, et al. Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology. 2010;53:344–53.
20. Ghedin E, Claverie J-MM. Mimivirus relatives in the Sargasso sea. Virol J. 2005;2:1–62.
21. Zhang W, Zhou J, Liu T, Yu Y, Pan Y, Yan S, et al. Four novel algal virus genomes discovered from Yellowstone Lake metagenomes. Sci Rep. 2015;5:1–13.
22. Kerepesi C, Grolmusz V. The "Giant Virus Finder" discovers an abundance of giant viruses in the Antarctic dry valleys. Arch Virol. 2017;162:1–6.
23. 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;56:386–94.
24. Popgeorgiev N, Colson P, Thuret I, Chiarioni J, Gallian P, Raoult D, et al. Marseillevirus prevalence in multitransfused patients suggests blood transmission. J Clin Virol. 2013;58:722–5.
25. Saadi H, DGI R, Colson P, Aherfi S, Minodier P, Pagnier I, et al. Shan virus: a new mimivirus isolated from the stool of a tunisian patient with pneumonia. Intervirology. 2013;56:424–9.
26. Dornas FP, Rodrigues FP, PVM B, LCF S, PCP F, Bonjardim CA, et al. Mimivirus circulation among wild and domestic mammals, Amazon Region, Brazil. Emerg Infect Dis. 2014;20:469–72.
27. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, et al. Faustovirus, an Asfarvirus-related new lineage of giant viruses infecting amoebae. J Virol. 2015;89:6585–94.
28. Colson P, Aherfi S, La Scola B, Raoult D. The role of giant viruses of amoebas in humans. Curr Opin Microbiol. 2016;31:199–208.
29. Aherfi S, Colson P, Raoult D. Marseillevirus in the pharynx of a patient with neurologic disorders. Emerg Infect Dis. 2016;22:2008–10.
30. Aherfi S, Colson P, Audoly G, Nappez C, Xerri L, Valensi A, et al. Marseillevirus in lymphoma: a giant in the lymph node. Lancet Infect Dis. 2016;1:225–34.
31. Popgeorgiev N, Michel G, Lepidi H, Raoult D, Desnues C. Marseillevirus adenitis in an 11-month-old child. J Clin Microbiol. 2013;51:4102–5.
32. Sambrook J, Russell DW. Molecular Cloning-Sambrook & Russel. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.
33. Reed LJ, Muench H. A simple method of estimating fifty per cent endpoints. Am J Hyg. 1938;27:493–7.
34. Campos RK, Boratto PV, Assis FL, Aguiar ER, Silva LC, Albarnaz JD, et al. Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. Virol J. 2014;11:95.
35. Andrade KR, Boratto PPVM, Rodrigues FP, LCF S, Dornas FP, Pilotto MR, et al. Oysters as hot spots for mimivirus isolation. Arch Virol. 2014;160:477–82.
36. Assis FL, Bajrai L, Abrahao JS, Kroon EG, Dornas FP, Andrade KR, et al. Pan-genome analysis of Brazilian lineage a amoebal mimiviruses. Viruses. 2015;7:3483–99.
37. dos Santos RN, Campos FS, Medeiros de Albuquerque NR, Finoketti F, Côrrea RA, Cano-Ortiz L, et al. A new marseillevirus isolated in Southern Brazil from Limnoperna fortunei. Sci Rep. 2016;6:35237.
38. Dornas FP, Assis FL, Aherfi S, Arantes T, Abrahão JS, Colson P, et al. A Brazilian marseillevirus is the founding member of a lineage in family marseilleviridae. Viruses. 2016;8:76.
39. PVM B, Arantes TS, LCF S, Assis FL, Kroon EG, La Scola B, et al. Niemeyer virus: a new mimivirus group A isolate harboring a set of duplicated aminoacyl-tRNA synthetase genes. Front Microbiol. 2015;6:1–11.
40. Antwerpen MH, Georgi E, Zoeller L, Woelfel R, Stoecker K, Scheid P. Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. Genome Announc. 2015;3:1–2.
41. Philippe N, Legendre M, Doutre G, Coute Y, Poirot O, Lescot M, et al. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science. 2013;341:281–6.
42. Scheid P. A strange endocytobiont revealed as largest virus. Curr Opin Microbiol. 2016;31:58–62.
43. Dutra LAL, de Freitas Almeida GM, Oliveira GP, Abrahão JS, Kroon EG, de Trindade GS. Molecular evidence of Orthopoxvirus DNA in capybara (Hydrochoerus hydrochaeris) stool samples. Arch Virol. 2017;162:439–48.
44. JYB K, Robert S, Reteno DG, Andreani J, Raoult D, La Scola B. High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. Front Microbiol. 2016;7:1–9.
45. Dare RK, Chittaganpitch M, Erdman DD. Screening pneumonia patients for Mimivirus1. Emerg Infect Dis. 2008;14:465–7.
46. Saadi H, Pagnier I, Colson P, Cherif JK, Beji M, Boughalmi M, et al. First isolation of Mimivirus in a patient with pneumonia. Clin Infect Dis. 2013;57:127–34.
47. Kerepesi C, Grolmusz V. Giant viruses of the Kutch Desert. Arch Virol. 2016; 161:721–4.
48. Santini S, Jeudy S, Bartoli J, Poirot O, Lescot M, Abergel C, et al. Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. Proc Natl Acad Sci. 2013;110:10800–5.
49. Yau S, Lauro FM, MZ DM, Brown MV, Thomas T, Raftery MJ, et al. Virophage control of antarctic algal host-virus dynamics. Proc Natl Acad Sci. 2011;108: 6163–8.

Additional file 1: Table S1: Identification and locations of viruses isolated.

Collections	Samples	Passage of isolation	Real Time PCR							Hemacolor Staining	TEM	SEM	Amoebas of isolation	Collection site	Coordinates	Date
Serra do Cipó			Mimiviridae				Pandoravirus	Marseillevirus	Cedratvirus							
			Lineage A	Llineage B	Lineage C	All lineages	Major Capsid Protein	Major Capsid Protein	Major Capsid Protein							
60	water	2°	neg	pos	neg	neg	neg	neg	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.840' W 43° 36.5911'	2015
46	water	2°	neg	pos	neg	neg	neg	neg	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
1	soil	3°	neg	pos	neg	neg	neg	neg	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
40	water	3°	neg	pos	neg	neg	neg	neg	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.756' W 43° 37.074'	2015
5	water	2°	neg	pos	neg	pos	neg	neg	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
59	water	3°	neg	pos	neg	neg	neg	pos	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
47	water	3°	neg	pos	neg	neg	neg	pos	ND	ND	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
12	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
3	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
22	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
32	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.756' W 43° 37.014'	2015
45	soil	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.756' W 43° 37.014'	2015
25	soil	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.756' W 43° 37.014'	2015
57	soil	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
29	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	mimivirus-like	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 22.595' W 43° 35.731'	2015
11	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
48	water	3°	neg	pos	neg	neg	neg	pos	ND	mimivirus-like	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015
44	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. polyphaga*	Serra do Cipó, MG,Brazil	S 19° 20.893' W 43° 36.555'	2015

Additional file 1: Table S1: Identification and locations of viruses isolated.

Collections	Samples	Passage of isolation	Real Time PCR							Hemacolor Staining	TEM	SEM	Amoebas of isolation	Collection site	Coordinates	Date
Sewage creeks			Mimiviridae				Pandoravirus	Marseillevirus	Cedratvirus							
			Lineage A	Llineage B	Lineage C	All lineages	Major Capsid Protein	Major Capsid Protein	Major Capsid Protein							
4.5	Sewage	1°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Sarandi creek, Belo Horizonte, MG, Brazil	S 19° 86.084' W 43° 99727'	2016
4.2	Sewage	2°	neg	neg	neg	neg	neg	neg	ND	ND	ND	mimivirus-like	*A. castellanii*	Sarandi creek, Belo Horizonte, MG, Brazil	S 19° 86.084' W 43° 99727'	2016
5.9	Sewage	1°	neg	neg	neg	neg	neg	neg	ND	ND	ND	mimivirus-like	*A. castellanii*	Tijuco creek, Belo Horizonte ,MG, Brazil	S 19° 85978' W 43° 98141'	2016
8.5	Sewage	1°	neg	ND	ND	neg	neg	neg	ND	ND	ND	pandoravirus-like	*A. castellanii*	Mergulhão creek, Belo Horizonte, MG, Brazil	S 19°86.478' W 43° 97618'	2016
9.2	Sewage	2°	neg	ND	ND	neg	pos	neg	ND	ND	ND	pandoravirus-like	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19° 86.084' W 43° 99727'	2016
9.10	Sewage	2°	neg	ND	ND	neg	neg	neg	ND	ND	ND	marseillevirus-lke	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2016
6.10	Sewage	2°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Tijuco creek, Belo Horizonte MG, Brazil	S 19° 85978' W 43° 98141'	2016
1.7	Sewage	2°	pos	ND	ND	pos	neg	neg	ND	ND	ND	mimivirus-like	*A. castellanii*	Sewage treatment Belo Horizonte, MG, Brazil	S 19° 86.084' W 43° 99723'	2016
2.9	Sewage	2°	pos	ND	ND	neg	neg	neg	ND	ND	ND	ND	*A. castellanii*	Conflux of Sarandi and Ressaca creek, Belo Horizonte, MG, Brazil	S 19° 86.084' W 43° 99727'	2016
8.7	Sewage	2°	pos	ND	ND	neg	neg	neg	ND	ND	ND	pandoravirus-like	*A. castellanii*	Mergulhão creek, Belo Horizonte, MG, Brazil	S 19°86.478' W 43° 97618'	2016

Additional file 1: Table S1: Identification and locations of viruses isolated.

Collections	Samples	Passage of isolation	Real Time PCR							Hemacolor Staining	TEM	SEM	Amoebas of isolation	Collection site	Coordinates	Date
Sewage creeks			Mimiviridae				Pandoravirus	Marseillevirus	Cedratvirus							
			Lineage A	Llineage B	Lineage C	All lineages	Major Capsid Protein	Major Capsid Protein	Major Capsid Protein							
17.1	Sewage	3°	neg	neg	pos	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17.2	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	ND	mimivirus-like	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17.5	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17.9	Sewage	3°	neg	neg	neg	neg	neg	neg	ND	mimivirus-like	ND	mimivirus-like	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.1	Sewage	2°	neg	neg	neg	neg	neg	neg	ND	mimivirus-like	ND	mimivirus-like	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.2	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.4	Sewage	3°	pos	ND	ND	pos	neg	neg	ND	ND	mimivirus-like	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.6	Sewage	3°	pos	ND	ND	neg	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.7	Sewage	2°	neg	neg	neg	neg	neg	neg	ND	mimivirus-like	ND	mimivirus-like	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.8	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017

Additional file 1: Table S1: Identification and locations of viruses isolated.

Collections	Samples	Passage of isolation	Real Time PCR							Hemacolor Staining	TEM	SEM	Amoebas of isolation	Collection site	Coordinates	Date
Sewage creeks			Mimiviridae				Pandoravirus	Marseillevirus	Cedratvirus							
			Lineage A	Llineage B	Lineage C	All lineages	Major Capsid Protein	Major Capsid Protein	Major Capsid Protein							
17b.9	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	mimivirus-like	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
17b.10	Sewage	2°	neg	neg	pos	neg	neg	neg	ND	mimivirus-like	mimivirus-like	ND	*A. castellanii*	Bom Jesus creek, Belo Horizonte, MG, Brazil	S 19°86084' W 43° 99727'	2017
Farm Sewage																
18.1	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	Itaúna, MG, Brasil	S 20°4'56.2044"	2017
18.4	Sewage	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	marseillevirus -like	ND	*A. castellanii*	Itaúna, MG, Brasil	S 20°4'56.2044"	2017
18.5	Sewage	2°	neg	ND	ND	neg	neg	neg	pos	cedratvirus-like	ND	cedratvirus-like	*A. castellanii*	Itaúna, MG, Brasil	S 20°4'56.2044"	2017
18.10	Sewage	3°	pos	ND	ND	pos	neg	neg	ND		ND	ND	*A. castellanii*	Itaúna, MG, Brasil	S 20°4'56.2044"	2017
Water treatment station																
19.1	water	3°	neg	neg	neg	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 9°58'36.0876" W 3°56'529368"	2017
19.2	water	3°	neg	neg	neg	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.4	water	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	mimivirus-like	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.5	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.6	water	2°	neg	neg	neg	neg	neg	neg	ND	ND	ND	mimivirus-like	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017

Additional file 1: Table S1: Identification and locations of viruses isolated.

Collections	Samples	Passage of isolation	Real Time PCR							Hemacolor Staining	TEM	SEM	Amoebas of isolation	Collection site	Coordinates	Date
Water treatment station			Mimiviridae				Pandoravirus	Marseillevirus	Cedratvirus							
			Lineage A	Llineage B	Lineage C	All lineages	Major Capsid Protein	Major Capsid Protein	Major Capsid Protein							
19.7	water	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.8	water	3°	neg	neg	neg	pos	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.9	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
19.10	water	3°	neg	neg	neg	pos	neg	neg	ND		mimivirus-like	ND	*A. castellanii*	COPASA, Belo Horizonte, MG, Brazil	S 19°58'36.0876" W 43°56'529368"	2017
Antarctic																
PP	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Punta Plaza, Antarctica	S 62° 06. 169' W 58° 21.375'	2014
PU	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	mimivirus-like	ND	*A. castellanii*	PuntaUllman, Antarctica	S 62° 05. 090' W 58° 20.592'	2014
YP	water	3°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Yellow Point, Antarctica	S 62° 05. 090' W 58° 20.592'	2014
ART	water	2°	pos	ND	ND	pos	neg	neg	ND	ND	mimivirus-like	ND	*A. castellanii*	Arctowski, Antarctica	S 62° 09. 246' W 58° 27.207'	2014
Others																
CAP10	Capybara feces	1°	pos	ND	ND	pos	neg	neg	ND	ND	ND	ND	*A. castellanii*	Serra do Cipó, MG,Brazil	S 19°20.457' W 43°37.008'	2012
MG collection	soil	3°	neg	ND	ND	neg	neg	neg	ND	mimivirus-like	ND	ND	*A. castellanii*	MG, Brazil	S 21° 64.617' W 45°44.003'	2014
MG collection	soil	1°	neg	ND	ND	neg	neg	neg	ND	ND	mimivirus-like	ND	*A. castellanii*	MG, Brazil	S 21° 23.437' W 42°80.067'	2014
Pantanal	soil	1°	neg	ND	ND	neg	neg	neg	ND	ND	tupanvírus	Unusual morphology	*A. castellanii*	Pantanal, MT, Brazil	-	2015

4.2. Artigo #2: Filling gaps about mimivirus entry, uncoating and morphogenesis

A investigação do ciclo de multiplicação de um vírus é essencial para uma melhor compreensão da biologia viral. Desde a descoberta dos mimivírus, um interesse crescente pelo estudo deste vírus se iniciou, entretanto, muitas etapas do ciclo de multiplicação destes vírus permanecem pouco compreendidas. Neste trabalho, novas informações a respeito do ciclo de multiplicação dos mimivírus são fornecidas, a partir de ensaios biológicos e de uma extensa análise de imagens de microscopia eletrônica destes vírus com enfoque nas etapas de penetração, desnudamento e morfogênese. Ensaios utilizando citocalasina, um inibidor de fagocitose, impactaram negativamente na incorporação dos vírus pela célula hospedeira *Acanthamoeba castellanii*, com consequente diminuição na multiplicação destes vírus. O tratamento com bafilomicina, um inibidor de acidificação de fagossomo, também impactou na multiplicação destes vírus. Aliado a isso, inúmeras imagens de microscopia eletrônica de transmissão mostraram a ocorrência da fusão do fagossomo contendo o vírus com o lisossomo com posterior início do desnudamento viral, sugerindo que a acidificação do fagossomo é importante nesta etapa do ciclo de multiplicação viral. No que se refere a morfogênese o capsídeo destes vírus, essa estrutura parece ser montada à partir estruturas lamelares crescentes visualizadas no citoplasma da célula hospedeira. Apesar de ter sido proposto anteriormente que a aquisição de genoma pelas partículas em formação ocorrer antes da aquisição de fibrilas, este estudo demonstra que a aquisição das fibrilas e do genoma pode ocorrer de forma simultânea. Além disso, foi sugerida a existência de uma área específica na periferia das fábricas virais dos mimivírus onde as partículas adquirem as fibrilas. Esta região foi denominada de área de aquisição de fibrilas. Por fim, foi evidenciado que partículas defectivas são formadas mesmo na ausência de virófagos. A partir destes novos achados um modelo mais atualizado do ciclo de multiplicação dos mimivírus foi proposto.

Este artigo foi publicado no periódico ***Journal of Virology*** em setembro de 2017

Após a publicação desse artigo nós também investigamos como é a área de aquisição de fibrilas em amostras de mimivírus que apresentam um número de fibrilas reduzido como nas amostras M4 e um novo isolado da linhagem B, denominado Borely moumouvirus. A amostra de mimivírus M4 foi obtido após 150 passagens em cultura de *A. castellanii,* resultando em uma deleção de 16% no seu genoma e uma redução drástica no número de fibrilas na partícula (BOYER *et al*., 2011). O isolado Borely moumouvirus foi obtido no estudo de prospecção viral apresentado nesta tese e foi isolado de amostra de água coletada na Serra do Cipó em Minas Gerais. Após a obtenção de imagens de MET do novo vírus isolado notamos que este vírus apresentava uma quantidade de fibrilas reduzidas quando comparado ao APMV (Figura 2). As imagens de MET mostram que estes vírus também apresentam na periferia de suas respectivas FV uma região de característica semelhante a observada para o APMV.

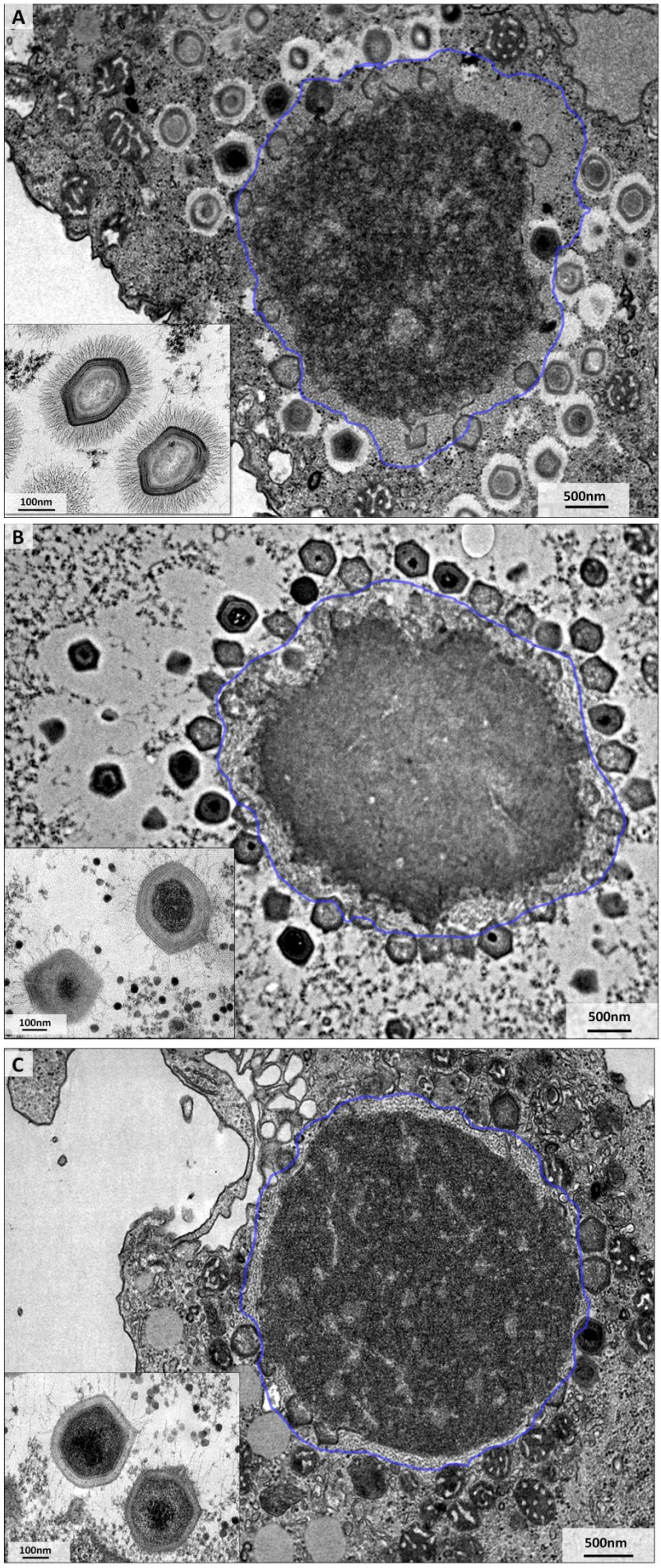

A
100nm
500nm
B
100nm
500nm
C
100nm
500nm

Figura 2: Área de aquisição de fibrilas (AAF) formada durante a morfogênese de diferentes amostras de mimivírus. Célula de *A. castellanii* infectada com diferentes amostras de mimivírus em um ciclo assincrônico. A AAF está destacada em azul. Em detalhe está a estrutura da partícula viral. A) APMV; B) Borely moumouvirus; C) mimivírus M4. Fontes: Banco de imagens do Gepvig.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Filling Knowledge Gaps for Mimivirus Entry, Uncoating, and Morphogenesis

Ana Cláudia dos Santos Pereira Andrade,[a] Rodrigo Araújo Lima Rodrigues,[a] Graziele Pereira Oliveira,[a] Kétyllen Reis Andrade,[b] Cláudio Antônio Bonjardim,[a] Bernard La Scola,[c] Erna Geessien Kroon,[a] Jônatas Santos Abrahão[a]

Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[a]; Laboratório de Imunopatologia de Doenças Virais, Centro de Pesquisas René Rachou—Fiocruz, Minas Gerais, Brazil[b]; URMITE, CNRS UMR 6236, IRD 3R198, Aix Marseille Universite, Marseille, France[c]

ABSTRACT Since the discovery of mimivirus, its unusual structural and genomic features have raised great interest in the study of its biology; however, many aspects concerning its replication cycle remain uncertain. In this study, extensive analyses of electron microscope images, as well as biological assay results, shed light on unclear points concerning the mimivirus replication cycle. We found that treatment with cytochalasin, a phagocytosis inhibitor, negatively impacted the incorporation of mimivirus particles by *Acanthamoeba castellanii*, causing a negative effect on viral growth in amoeba monolayers. Treatment of amoebas with bafilomicin significantly impacted mimivirus uncoating and replication. In conjunction with microscopic analyses, these data suggest that mimiviruses indeed depend on phagocytosis for entry into amoebas, and particle uncoating (and stargate opening) appears to be dependent on phagosome acidification. In-depth analyses of particle morphogenesis suggest that the mimivirus capsids are assembled from growing lamellar structures. Despite proposals from previous studies that genome acquisition occurs before the acquisition of fibrils, our results clearly demonstrate that the genome and fibrils can be acquired simultaneously. Our data suggest the existence of a specific area surrounding the core of the viral factory where particles acquire the surface fibrils. Furthermore, we reinforce the concept that defective particles can be formed even in the absence of virophages. Our work provides new information about unexplored steps in the life cycle of mimivirus.

IMPORTANCE Investigating the viral life cycle is essential to a better understanding of virus biology. The combination of biological assays and microscopic images allows a clear view of the biological features of viruses. Since the discovery of mimivirus, many studies have been conducted to characterize its replication cycle, but many knowledge gaps remain to be filled. In this study, we conducted a new examination of the replication cycle of mimivirus and provide new evidence concerning some stages of the cycle which were previously unclear, mainly entry, uncoating, and morphogenesis. Furthermore, we demonstrate that atypical virion morphologies can occur even in the absence of virophages. Our results, along with previous data, allow us to present an ultimate model for the mimivirus replication cycle.

KEYWORDS mimivirus, electron microscopy, replication cycle, phagocytosis, fibril acquisition area

The giant Acanthamoeba polyphaga mimivirus (APMV), which is associated with amoebas of the *Acanthamoeba* genus, was isolated in 2003 and astonished the scientific community with unusual structural and genomic features within the virosphere (1, 2). In subsequent years, several mimivirus-like viruses were uncovered in different parts of the world, thus expanding the *Mimiviridae* family, especially the

Received 2 August 2017 **Accepted** 23 August 2017

Accepted manuscript posted online 6 September 2017

Citation Andrade ACDSP, Rodrigues RAL, Oliveira GP, Andrade KR, Bonjardim CA, La Scola B, Kroon EG, Abrahão JS. 2017. Filling knowledge gaps for mimivirus entry, uncoating, and morphogenesis. J Virol 91:e01335-17. https://doi.org/10.1128/JVI.01335-17.

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., R.A.L.R., and G.P.O. contributed equally to this work.

Mimivirus genus (3–7). These viruses have some genetic differences which define three distinct lineages (A, B, and C), but they are structurally similar.

Mimiviruses are composed of a particle with a pseudoicosahedral symmetry, 750-nm diameter, and a genome of double-stranded DNA of approximately 1.2 Mb (2, 8). They have a capsid of 500 nm that is formed of multiple protein layers and a lipid membrane surrounding an inner proteinaceous core, which contains the genome. A star-shaped projection is present on the capsid, from which the viral genome is released into the host cytoplasm; this is referred to as the stargate and represents a unique feature of mimiviruses (9). In addition, a dense layer of 125-nm-long glycoproteic fibers covers the viral surface and is important for viral attachment to different organisms, including amoeboid hosts (10).

Given the large size of mimiviruses, it has been proposed and largely accepted that the replication cycle of these viruses begins with phagocytosis of viral particles by *Acanthamoeba* spp. cells (11, 12). Information is lacking concerning the set of events that occurs between virion entry and stargate opening, but there is evidence that after opening of the stargate, the virion inner lipid membrane merges with that of the entry vesicle and releases the viral seed into the amoeba's cytoplasm. A typical viral eclipse phase is then established, during which viral particles are not visible in the cell (11, 13). The viral seed releases the DNA, promoting a reorganization of the host cytoplasm with further formation of viral factories (VF), where the virus genome is replicated and transcribed and new particles are assembled (13, 14). Based on atomic force microscopy, Kuznetsov and colleagues proposed a model for morphogenesis of mimivirus particles (15). In the model, capsid assembly occurs on the VF surface in a temporal fashion; assembly is initiated by the formation of the stargate portal, followed by thickening of the protein layer in its immediate vicinity. Before capsid formation is complete, the capsids are filled with DNA and other macromolecules through a stargate distal portal. Once the genome is enclosed within the capsid, an integument protein layer attaches to the capsid, on which a coating of fibrils then adheres to the entire viral surface (15). The involvement of the nucleus during the replication of these viruses is still unclear. Suzan-Monti and colleagues (11) suggested that APMV replication is nucleocytoplasmic, although other studies have argued that replication occurs exclusively in the host cell cytoplasm (13).

Mimiviruses have been studied for several years; nevertheless, some aspects regarding their biology remain unclear. Based on biological assays in conjunction with extensive electron microscopic analyses, now we are able to shed new light on the mimivirus replication cycle.

RESULTS

Biological assays suggest mimivirus entry into amoebas by phagocytosis. Although mimiviruses have been studied for more than a decade, some aspects of their replication cycle remain uncertain, such as their entry into natural hosts. Due to its large size (~700 nm), it has been proposed and widely accepted that mimivirus enters *Acanthamoeba* cells by phagocytosis (11, 13). However, to our knowledge there has been no biological demonstration of this process. Here, we confirm this hypothesis by using experimental data. Pretreatment of cells with cytochalasin D significantly decreased the titers and numbers of particles incorporated into cells infected with APMV (Fig. 1A and B). *Marseillevirus marseillevirus* (MsV) titers and particle incorporation were not reduced after this treatment, since isolated purified particles exploit the endocytic pathway (16). Pretreatment with bafilomycin also resulted in a significant reduction in APMV and MsV titers (Fig. 1C). Bafilomycin is a specific inhibitor of vacuolar-type H^+/ATPase that prevents phagosome and endosome acidification (17). Therefore, these results demonstrated that APMV replication is dependent on the acidification of phagosomes for the uncoating step. The acidification of the entry compartment is also important for MsV (16). These results have been corroborated by several transmission electron microscope (TEM) images that show the occurrence of phagosome-lysosome fusion, subsequent stargate opening, and release of the viral seed (Fig. 1D to L). Taken

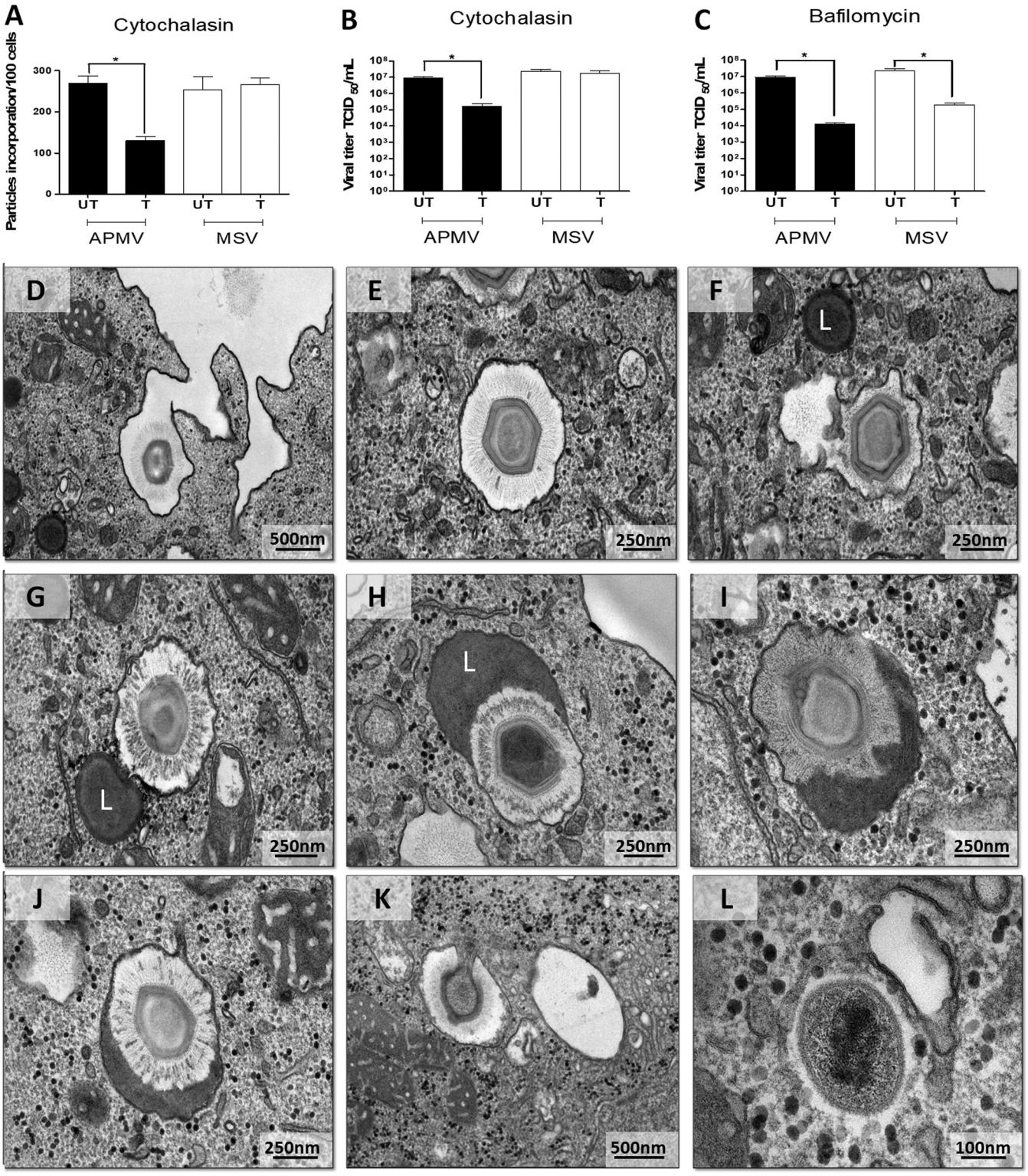


FIG 1 Mimivirus entry into the host cell by phagocytosis and uncoating depends on acidification of the phagosome. (A to C) The impact of cytochalasin D and bafilomycin on mimivirus replication. Treatment with cytochalasin D and bafilomycin decreased the mimivirus titers and particle incorporation. MsV was used as a control, and its titer was not reduced after the treatment with cytochalasin D but was reduced with treatment with bafilomycin. (D to L) Transmission electron microscopy images of mimivirus particles entering cells by phagocytosis (D to F), the phagosome-lysosome fusion (F to J), stargate opening (K), and release of the viral seed (L). The inhibitory assays were performed in triplicate. Error bars indicate standard deviations. *, $P < 0.05$ (two-tailed Student's *t* test). UN, untreated cells; T, treated cells; L, lysosome.

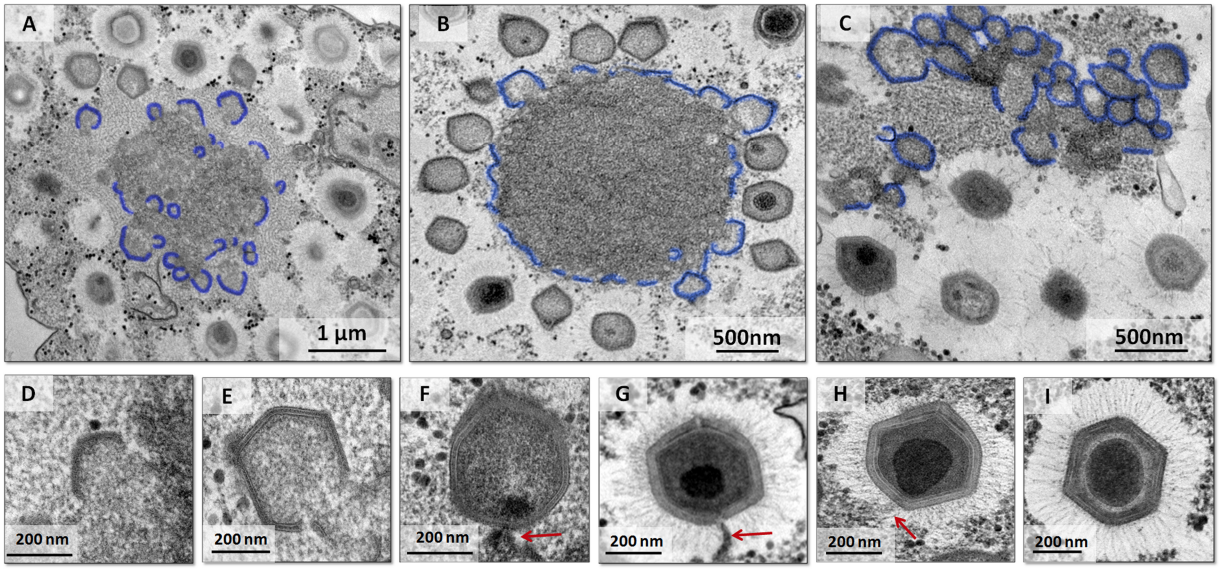


FIG 2 Growing lamellar structures seem to be important to mimivirus capsid assembly. (A to C) Growing lamellar structures with different sizes in the viral factory, demonstrating the viral crescent-like structures (in blue). The smaller blocks are observed in the interiors of the viral factories, and more complete particles are shown in the periphery of the factory. (D to I) Transmission electron microscopy images showing stages of mimivirus particle formation, as the particle grows in thickness and complexity. Images show the viral crescent (D), particles without a genome (E), fibrils (E and F), particles acquiring genome (red arrows) and fibrils (F to H), and finally complete mimivirus particle formation (I).

together, these results support the view that APMV entry into *Acanthamoeba* cells may occur via phagocytosis.

Capsids are assembled from growing lamellar structures. After mimivirus enters amoeba cells and delivers the viral seed, early DNA replication and transcription appear to take place exclusively in the host cytoplasm, moments before the formation of a mature VF (13). Once a mature VF is formed, the morphogenesis step is initiated. It has been proposed that the formation of the mimivirus capsid begins with the formation of the stargate, followed by thickening of the protein layer in its immediate vicinity (15). However, it is not yet clear how expansion of this protein layer occurs. Through analyses of TEM images of VF at different time points postinfection, we observed the presence of several lamellar structures preceding capsid formation, with different sizes in the inner part of the VF (Fig. 2A to C and 3A and B). Mutsafi and colleagues in 2013 showed the formation of multivesicular bodies and membrane sheets from the endoplasmic reticulum (ER) in the outermost zone of the mimivirus VF, the membrane assembly zone (18). Our TEM images provide evidence for the formation of lamellar structures, not only on the periphery of the VF but also in its interior (Fig. 2A to C and 3A and B), which is in accordance with the data of Mutsafi and colleagues. It is noteworthy that after assembly of the complete particle in the later stages of infection (~7 h postinfection), the crescents were no longer observed. These lamellar structures, which began in a way analogous to the crescents described for poxviruses and marseilleviruses, increased in complexity until reaching the VF periphery and moving to the next stages of virion morphogenesis (Fig. 2D to I): incorporation of DNA and surface fibrils. In contrast to lamellar capsid-forming structures, the main component of viral factories is a globular aspect, likely formed by DNA, membranes, and enzymes (Fig. 3A and B).

Dynamic acquisition of surface fibrils: the fibril acquisition area (FAA). The mature mimivirus particle is covered by fibrils of lengths ranging from 125 nm to 140 nm (8). A previous study based on atomic force microscopy proposed that the surface fibers are acquired by viral capsids when they pass sequentially through a membrane embedded with integument protein and then through a membrane containing surface fibers (15). In this model, layers of integument and fibrils are acquired

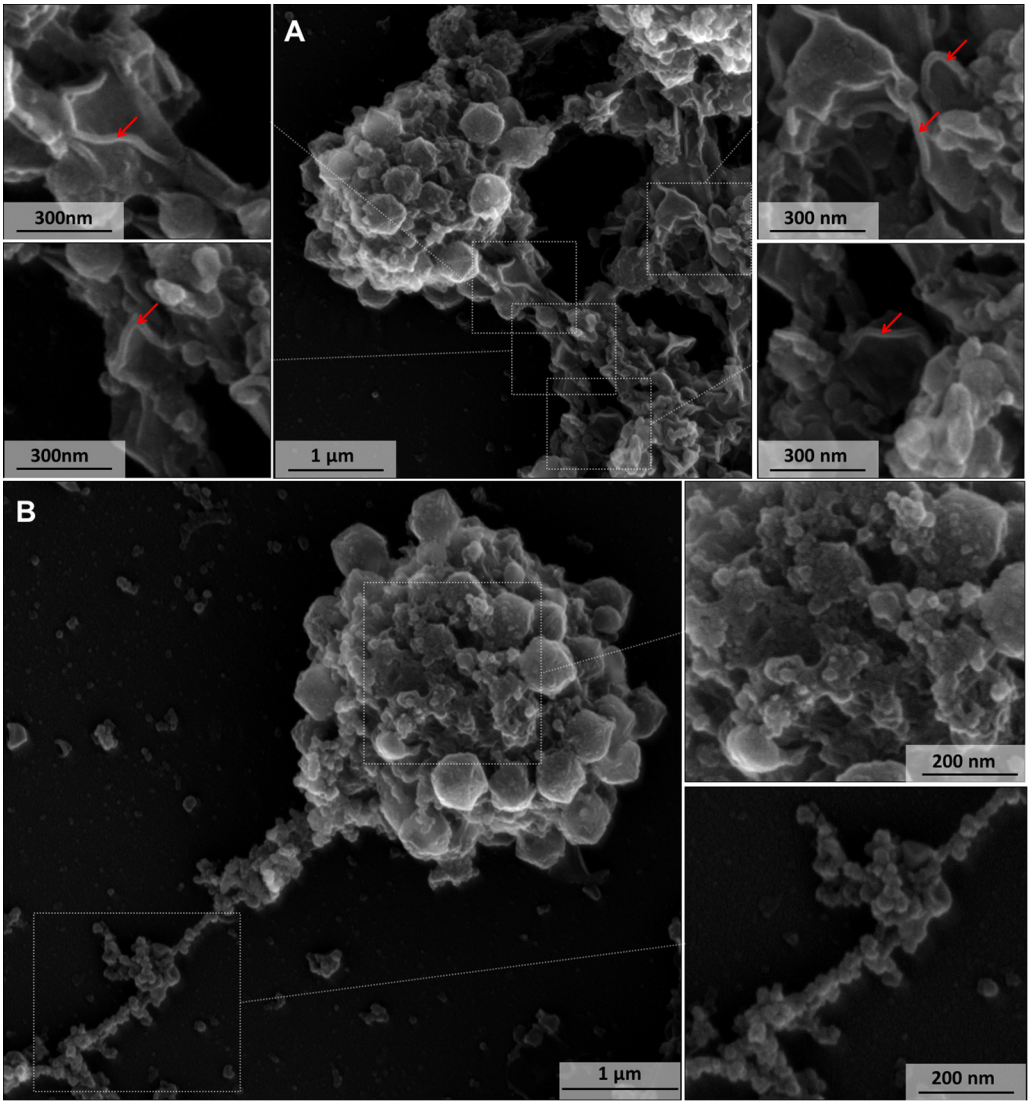


FIG 3 Scanning microscopy of a disrupted mimivirus viral factory. (A) Lamellar structures related to capsid morphogenesis. The red arrows indicate the lamelar structures around the viral factory. (B) In contrast to lamellar capsid-forming structures, the main component of VF presents a globular aspect, likely formed by DNA, membranes, and enzymes.

as an envelope involving the capsid. The same study also suggested that the integument protein layer is acquired near the VF but the fibril layer is acquired near the cellular periphery, and the source of these fibrils has not been demonstrated (15).

We noticed a very clear phenomenon regarding the dynamics of fibril acquisition, leading us to an overview of this step. In TEM images, it is possible to observe at the periphery of the VF a less-electron-dense region of an apparent fibrillar nature, which we named the fibril acquisition area (Fig. 4A and B). This arrangement is very clear and can be verified in almost all VF. During particle morphogenesis, newly formed capsids pass through this region where the particles acquire the fibrils. This was evident when we observed particles at the edge of the FAA, which had fibrils only in the upper portion, suggesting that this fibrillar region provides the fibrils for these particles (Fig. 5). In addition, the particles that had already passed through this region had fibrils all over their surface (Fig. 5). In the stages where we found few formed viral particles, the dimension of FAA was larger, and as long as the amount of newly formed particles increased, this region was reduced in size (Fig. 4A and B and 5). This phenomenon suggests that, as viral particles are formed, the fibrillar material that gives rise to fibrils

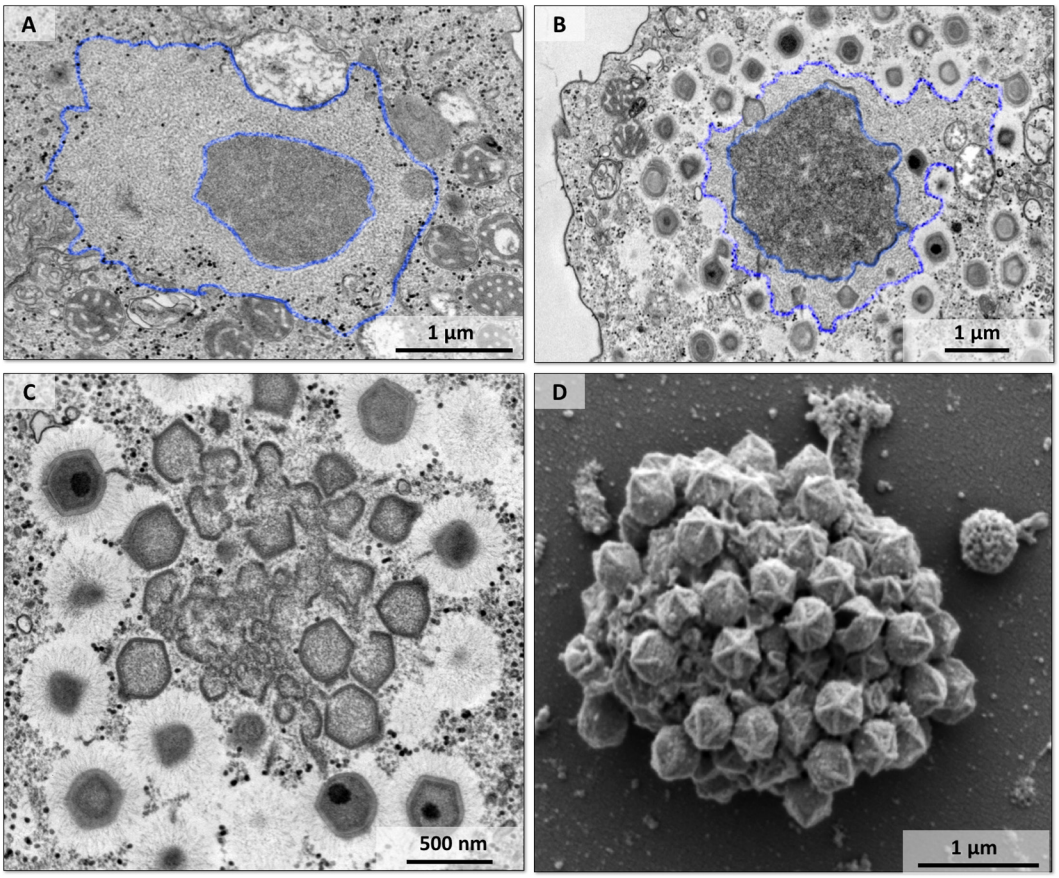


FIG 4 An area of fibril acquisition is present in the periphery of the viral factory. (A to C) Transmission electron microscopy images of the viral factory of mimivirus with a less-electron-dense area surrounding the periphery of the factory that decreased in size during particle morphogenesis. A fibril acquisition area is shown at early (A) and mature (B) stages of viral factory formation (highlighted by blue lines). (C) Mimivirus particles with fibrils in the periphery of the viral factory and particles without fibrils inside the factory before passing through the fibril acquisition area. (D) Scanning electron microscopy image of a viral factory with particles without fibrils (before fibril acquisition) and a prominent stargate.

is consumed. The TEM images of the VF showed the viral capsids forming blocks, growing in thickness and complexity as they migrated to the outermost part of the VF and acquiring fibrils when they passed through the FAA (Fig. 2 and 4C). In a scanning electron microscope (SEM) image, it was not possible to view the FAA, and the particles were being assembled before fibril acquisition (Fig. 4D). Notably, all particles were without fibrils and displayed a prominent stargate (Fig. 4D).

Together, the data presented here lead us to suggest that the fibrils are acquired in the FAA, not near the cellular periphery as previously suggested (15). This view was also supported by the presence of various viral particles with fibrils in the cell cytoplasm. In addition, we did not observe the acquisition of the fibril layer during passage through a membrane, as previously described (15). Our images suggest that fibrils are acquired by the capsid when they pass through a region presenting the preformed fibrils. The full composition of the FAA, as well as the mechanism(s) of fibril binding in the protein capsid, still require further investigation.

Simultaneous occurrence of genome incorporation and fibril acquisition. The processes of release and packaging of the viral genome in mimiviruses have been analyzed in previous studies (9, 13, 15). It has been demonstrated that genome delivery occurs through the stargate and packaging occurs through the opposite portal (9). This strategy is different from that of other icosahedral viruses, in which a single vertex-portal system plays a crucial role in both genome release and packaging (9, 19). The viral genome is acquired at the periphery of the VF during the morphogenesis step

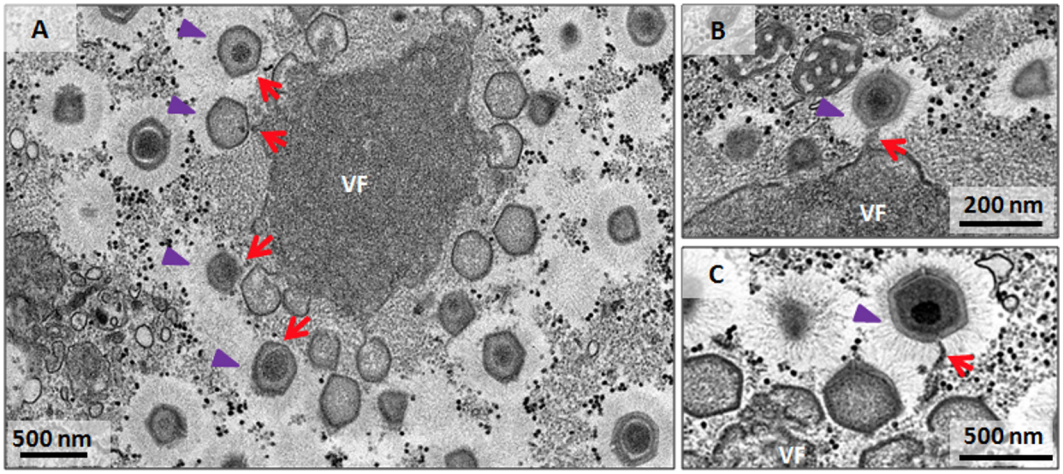


FIG 5 Fibril acquisition and genome incorporation can occur simultaneously. Transmission electron microscopy images of mature viral factories indicate genome incorporation (red arrows) simultaneous to the acquisition of fibrils (purple arrowheads). Viral particles are not fully covered by the fibrils (only half of each particle has fibrils), and the genome is still being packaged. The image shows that the viral genome is acquired at the periphery of the viral factory, at the FAA.

(11, 13). The particles that have already acquired a genome have a more-electron-dense region inside them, while those that have not yet acquired the genome do not exhibit this feature (Fig. 2E to H and 5). It was previously proposed that the formation of the capsid, acquisition of the genome, and acquisition of fibrils by mimivirus particles occur sequentially in this order (15). According to this model, the genome enters into the empty capsid, the nucleic acid condenses inside, and the entry portal is sealed, completing the assembly of the capsid. Subsequently, upper structural layers of integument and fibrils are acquired to form a mature particle (15).

Our data strongly suggest an alternative view of these events during mimivirus morphogenesis. In transverse sections of the VF, we clearly observed particles which were fully or partially covered by fibrils still acquiring the genome (Fig. 4C and 5). Further TEM images showed particles that had already passed through the fibrillar region and were therefore covered by fibrils and were acquiring the genome in the region opposite the stargate (Fig. 4C and 5). It is noteworthy that some viral particles were not fully covered by the fibrils (only half of the particle has fibrils), and the genome was still being packaged (Fig. 5). In contrast to what has been previously proposed, our data suggest that no sequential order of events occurs in the final steps of mimivirus morphogenesis concerning genome acquisition and fibril incorporation. It is possible that fibrils attach to some viral particles after full genome acquisition, but most of the images presented here lead us to support the hypothesis that the acquisition of the genome occurs concomitantly with the acquisition of the surface fibrils.

Viral particles with unusual morphologies. The particles of viral progeny were mostly well formed with a typical morphology. However, particles with unusual structures were also observed. TEM images showed unusual projections starting from the capsid or on its sides and from capsids without typical symmetry (Fig. 6A to C). In addition, an assembled empty capsid in the center of the VF was observed, an unexpected finding considering the particular stage of the virus replication cycle (Fig. 6D). Particles with an altered fibril arrangement were also identified, and SEM images showed particles with a nonuniform distribution of fibrils (Fig. 6E and F).

It has been shown that virophages affect the replication cycle of mimiviruses, decreasing amoeba lysis and generating particles with abnormal morphologies (20–22). Some particles showed capsid layers that asymetrically accumulated on the viral particle or harbored fibrils in only one part of the capsid. Sputnik was the first virophage described which infected a strain of mimivirus (20). Since then, other virophages have

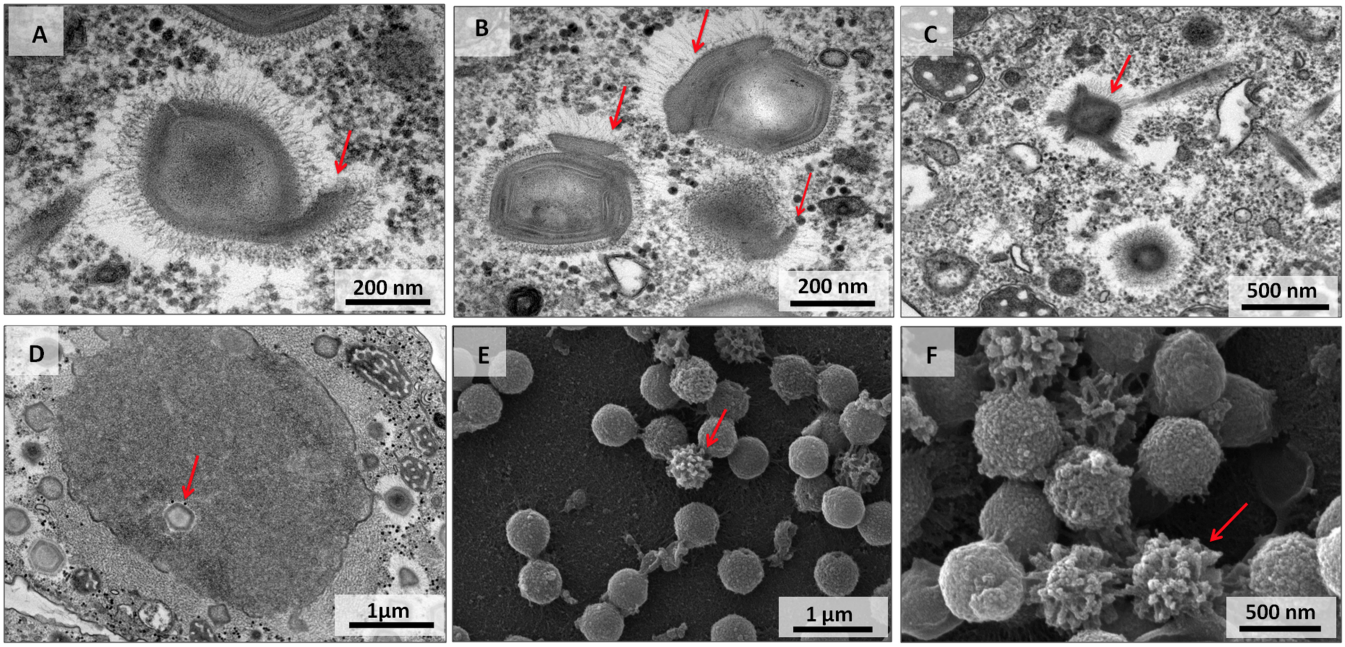


FIG 6 Defective particles in the mimivirus replication cycle in the absence of virophages. (A to D) Transmission electron microscopy images of mimivirus particles with atypical morphology (A to C) and an unusual localization of assembled empty capsid in the center of the viral factory (D). (E and F) Scanning electron microscopy images of particles with a nonuniform distribution of fibrils. Red arrows point to unusual features.

been described, thus reinforcing this new class of viruses (5, 22, 23). Virophages cannot replicate alone in amoebas, but they can replicate in the mimiviral VF. They have a small diameter (50 nm) and an icosahedral capsid, but they can be visualized in TEM images in VF and even within viral particles (20, 22). However, none of the images analyzed in this study exhibited signs of the presence of virophages, even in those images in which defective particles were found. To date, our APMV stocks have been tested by PCR for the presence of Sputniks and Zamilon, but we failed to amplify any virophage gene.

As previously demonstrated, different multiplicities of infection (MOI) used for mimivirus production resulted in distinct proportions of infectious particles in newly formed progeny, whereas the most efficacious way to produce infectious mimivirus particles is to use a low MOI (24). Although we followed this strategy during sample preparation for TEM analyses, multiple particles can infect the same cell and saturate the cellular machinery, thus generating defective particles. This is probably a usual phenomenon in nature, and the formation of defective particles, even in the absence of virophages, may be more frequent than previously thought (20, 24).

DISCUSSION

A better comprehension of the biology of a virus is achieved with basic investigations of its life cycle. Viruses are a highly heterogeneous group of organisms and present different replication strategies. In general, DNA viruses replicate in the host cell nucleus, while RNA viruses replicate within the cell cytoplasm. However, there are some exceptions, such as some members of the putative *Megavirales* order: poxviruses and iridoviruses, for instance (25). These double-stranded DNA viruses replicate in the host cytoplasm, wherein they establish a complex viral factory as an intracellular compartment in which both genome replication and viral morphogenesis occur (26). The giant mimiviruses share this same peculiarity. Studies have been conducting to better comprehend the replication cycle of these viruses, wherein models that explain this virus life cycle are constantly being updated (11, 13, 15, 18). In this study, we filled some gaps in the data concerning mimiviruses and introduced some data relating to entry, uncoating, and morphogenesis in the mimivirus replication cycle.

Mimiviruses have a diameter of approximately 700 nm and infect free-living amoe-

bas from the *Acanthamoeba* genus, which are protists that feed by phagocytosis (8, 27). The large size of mimiviruses associated with their host's life style, along with electron microscopy images, led to the assumption that these viruses enter host cells by phagocytosis, but until now there were no biological data to validate these assumptions. Inhibition of phagocytosis with cytochalasin D allowed us to observe a significant reduction in mimivirus entry, which was not observed with marseilleviruses (viruses of ~200-nm diameter) (Fig. 1A and B). With supported from TEM images, the data show that mimiviruses entered amoeba cells by phagocytosis, while marseilleviruses explored alternative pathways, as previously demonstrated (16). Furthermore, our data suggested the possible importance of phagosome acidification for mimivirus uncoating (Fig. 1C to L). In the presence of bafilomycin, viral replication was impaired. It is possible that a decrease in pH after formation of the phagolysosome is the triggering factor to the opening of the mimivirus stargate, thus allowing the release of viral seed into the host cytoplasm. Nevertheless, it remains uncertain which pH is ideal for triggering this process, or whether enzymes present in the lysosomes are essential for this process.

After release of the viral seed, the DNA replication and transcription take place in early VF (13). It is possible that nuclear factors are involved in this step, but further evidence is needed (28). Assembly of new viral particles takes place inside a mature VF. Previous studies have shown that membranes from the endoplasmic reticulum (ER) migrate toward the edge of the VF and act as scaffolding for the assembly of capsids, which would occur by acquisition of individual pentameric units in sequence (15). Our data corroborate this theory, demonstrating the existence and expansion of lamellar structures, in an way analogous to that of crescents described for poxviruses and marseilleviruses (Fig. 2 and 3) (16, 29). It is possible that membranes from the ER migrate into the VF, wherein they act as scaffolding for protein blocks. These structures increase in size by incorporation of proteins which move away from the core of the VF, until they pass through a fibrillar aspect area, which we named the fibril acquisition area, wherein the new viral particles acquire surface fibrils (Fig. 4 and 5). This hypothesis was strengthened when we systematically observed the reduction of the fibrillar area after release of viral particles from the VF, suggesting that a fibrillar matrix is continuously consumed as the particles pass through it. Curiously, fibril proteins were not detected in isolated VF analyzed via proteomic approaches (14). It is possible that a large fraction of that region was lost during the purification process of the factories, thus hampering the detection of fibrils by this method. Our hypothesis for the acquisition of fibrils by mimiviruses relies on a different or maybe complementary way to that proposed by Kuznetsov and colleagues, who proposed that a layer containing the fibrils involves the viral capsid in a region distant from the VF (15). Moreover, those authors proposed that genome acquisition occurs before the acquisition of fibrils. Our images clearly show that these events occur simultaneously (Fig. 4C and 5). It is possible that some particles receive the genome while the capsid is assembled and then the fibrils are attached, as previously proposed, but we believe that, in most cases, these events occur concomitantly.

The data presented in this study, along with previous published data based on atomic force microscopy, X-ray diffraction, and fluorescence led us to present an updated model of the replication cycle of mimiviruses (Fig. 7). Viruses enter amoeba cells by phagocytosis (0 h). After fusion of the phagosome with the lysosome (1 to 2 h), the stargate is opened and the viral seed is released (4 h). An early VF is established and viral proteins are synthesized outside the factory (4 to 6 h). In a mature VF, viral crescents increase in complexity and acquire the genome and fibrils simultaneously, as long as they move from the core of the VF to the FAA (8 h). Among the newly formed viral particles, some of them might not be infectious (defective particles), and these would exhibit morphological features similar to those in the presence of virophages (20, 22). Finally, the viral progeny is released by cell lysis. The comprehension of the biology of mimiviruses is improving due to a combination of biological and genetic data, along with increasingly illuminating images from different microscopic techniques. Notwithstanding current developments, some aspects regarding the replication

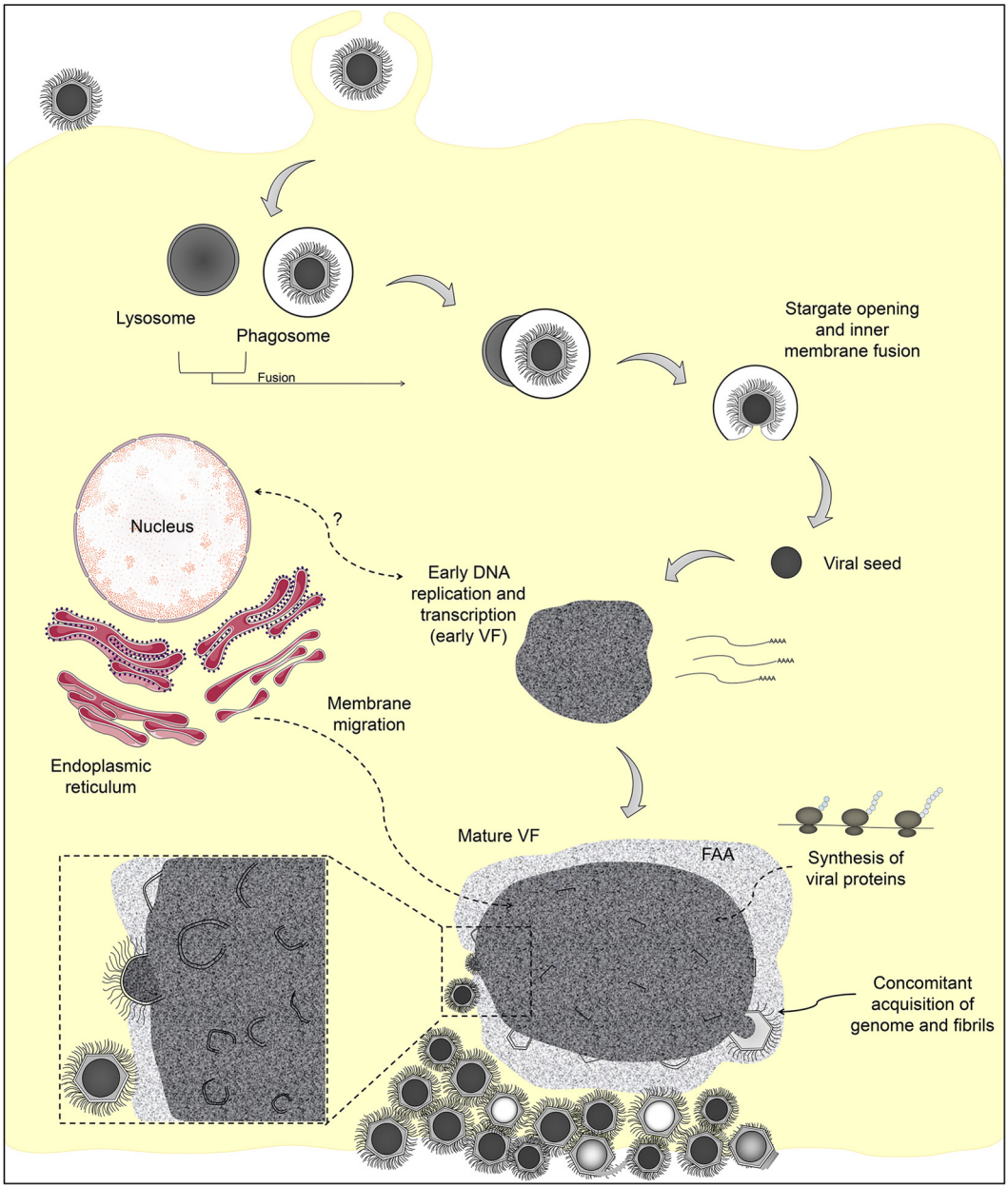


FIG 7 Representative scheme of the mimivirus replication cycle. Infectious mimivirus particles enter host cell by phagocytosis. Fusion of the phagosome with lysosome occurs and opening of the stargate is followed by release of the viral seed. An early viral factory is established, and viral proteins are synthesized outside the factory. It is uncertain whether the cell nucleus is involved in mimivirus genome replication. The viral crescents increase in thickness and complexity in the mature viral factory and might acquire genomic material and fibrils simultaneously. The particles move from the core of the viral factory to the FAA, where the particles incorporate the surface fibrils. Among the newly formed viral particles, some of them might not be infectious (defective particles), presenting atypical morphology. The viral progeny are released by cell lysis.

cycle of mimiviruses remain uncertain, and these mainly occur at the molecular level. Future studies will further enhance the model presented here and improve our understanding of the biology of these complex viruses.

MATERIALS AND METHODS

Cell culture and virus production, purification, and titration. *Acanthamoeba castellanii* (ATCC 30010) cells were cultivated in peptone-yeast extract-glucose (PYG) medium supplemented with 25 mg/ml amphotericin B (Fungizone; Cristalia, São Paulo, Brazil), 500 U/ml penicillin, and 50 mg/ml gentamicin (Schering-Plough, Brazil). A total of 7×10^6 cells were infected with APMV at an MOI of 0.01 and incubated at 32°C. After the appearance of a cytopathic effect, cells and supernatant were collected

and viruses were then purified through ultracentrifugation with a 25% sucrose cushion at 36,000 × *g* for 30 min. After purification, amoeba cells were infected with purified viruses at an MOI of 0.01. The viruses were serially diluted, and multiple replicate samples of each dilution were inoculated into *A. castellanii* (ATCC 30234) monolayers. After 72 to 96 h of incubation, the amoebas were analyzed to determine whether infection had taken place. Based on these data, the virus titers were determined using the endpoint method (30).

Transmission electron microscopy. The *Acanthamoeba castellanii* cells were infected as described in the previous section at an MOI of 0.01 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 then were analyzed under TEM (Spirit Biotwin FEI, 120 kV).

Scanning electron microscopy. The *Acanthamoeba 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 a 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 metallized with a 5-nm gold layer. The analyses were completed using scanning electronic microscopy (FEG Quanta 200 FEI).

Entry and uncoating assays. In the entry and uncoating experiments, we evaluated whether blocking of phagocytosis and phagosome acidification impacted APMV entry and replication. A total of 10^6 *A. castellanii* cells were treated with 2 μM cytochalasin, a phagocytosis inhibitor, in a total volume of 5 ml of PYG medium. After 1 h, supernatant was removed and cells were infected with APMV at an MOI of 5. As a negative control for the phagocytosis process, we used sonicated, purified particles of MsV under the same conditions described for APMV. It was previously demonstrated that isolated particles of MsV enter amoebas by endocytosis, not phagocytosis. Control groups of untreated, infected amoebas were also used. Two hours postinfection, the supernatant was collected to measure the remaining viral particles (nonphagocytized). The supernantant was also collected immediately after infection (0 h postinfection). The quantity of particles was calculated using flow cytometry, as previously described (31). The rate of particle incorporation was calculated, taking into consideration the variation in particle content of the supernatant between times 0 h and 2 h postinfection. An identical assay was performed in parallel, but after infection the monolayer was washed with PAS, fresh medium was added, and at 8 h postinfection amoebas were collected and titrated. The aim of this experiment was to estimate the impact on viral replication of the blocking of phagocytosis. Finally, to investigate whether blocking of cell acidification impacted APMV and MsV replication, cells were treated with 5 nM bafilomycin. Eight hours postinfection, cells were collected and titrated. All the experiments were performed three independent times, in duplicate. Graphs were constructed using GraphPad Prism version 7.00 for Windows (GraphPad Software).

ACKNOWLEDGMENTS

We thank our colleagues from Gepvig and the Laboratório de Vírus for their excellent technical support.

We also thank CNPq, CAPES, and FAPEMIG for scholarships and the Center of Microscopy of UFMG. J.S.A., C.A.B., and E.G.K. are CNPq researchers. J.S.A., E.G.K., and B.L.S. are members of a CAPES-COFECUB project.

REFERENCES

1. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie J-M, Raoult D. 2003. A giant virus in amoebae. Science 299:2033. https://doi.org/10.1126/science.1081867.
2. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, Scola B La, Suzan M, Claverie J-M. 2004. The 1.2-megabase genome sequence of mimivirus. Science 306:1344–1350. https://doi.org/10.1126/science.1101485.
3. Yoosuf N, Yutin N, Colson P, Shabalina SA, Pagnier I, Robert C, Azza S, Klose T, Wong J, Rossmann MG, La Scola B, Raoult D, Koonin EV. 2012. Related giant viruses in distant locations and different habitats: Acanthamoeba polyphaga moumouvirus represents a third lineage of the Mimiviridae that is close to the Megavirus lineage. Genome Biol Evol 4:1324–1330. https://doi.org/10.1093/gbe/evs109.
4. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie J-M. 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc Natl Acad Sci U S A 108:17486–17491. https://doi.org/10.1073/pnas.1110889108.
5. Campos RK, Boratto PV, Assis FL, Aguiar ER, Silva LC, Albarnaz JD, Dornas FP, Trindade GS, Ferreira PP, Marques JT, Robert C, Raoult D, Kroon EG, La Scola B, Abrahão JS. 2014. Samba virus: a novel mimivirus from a giant rainforest, the Brazilian Amazon. Virol J 11:95. https://doi.org/10.1186/1743-422X-11-95.
6. Dornas FP, Khalil JYB, Pagnier I, Raoult D, Abrahão 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.
7. Saadi H, Reteno DGI, Colson P, Aherfi S, Minodier P, Pagnier I, Raoult D, La Scola B. 2013. Shan virus: a new mimivirus isolated from the stool of a tunisian patient with pneumonia. Intervirology 56:424–429. https://doi.org/10.1159/000354564.
8. Xiao C, Kuznetso YG, Sun S, Hafenstein SL, Kostyuchenko VA, Chipman PR, Suzan-Monti M, Raoult D, McPherson A, Rossmann MG. 2009. Structural studies of the giant mimivirus. PLoS Biol 7:e92. https://doi.org/10.1371/journal.pbio.1000092.
9. Zauberman N, Mutsafi Y, Halevy D Ben, Shimoni E, Klein E, Xiao C, Sun S, Minsky A. 2008. Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus. PLoS Biol 6:1104–1114. https://doi.org/10.1371/journal.pbio.0060114.
10. Rodrigues RAL, dos Santos Silva LK, Dornas FP, de Oliveira DB, Magalhães TFF, Santos DA, Costa AO, de Macêdo Farias L, Magalhães PP,

Bonjardim CA, Kroon EG, La Scola B, Cortines JR, Abrahão JS. 2015. Mimivirus fibrils are important for viral attachment to the microbial world by a diverse glycoside interaction repertoire. J Virol 89: 11812–11819. https://doi.org/10.1128/JVI.01976-15.
11. Suzan-Monti M, La Scola B, Barrassi L, Espinosa L, Raoult D. 2007. Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga mimivirus. PLoS One 2:e328. https://doi.org/10.1371/journal.pone.0000328.
12. Abrahão JS, Dornas FP, Silva LC, Almeida GM, Boratto PV, Colson P, La Scola B, Kroon EG. 2014. Acanthamoeba polyphaga mimivirus and other giant viruses: an open field to outstanding discoveries. Virol J 11:120. https://doi.org/10.1186/1743-422X-11-120.
13. 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.
14. Fridmann-Sirkis Y, Milrot E, Mutsafi Y, Ben-Dor S, Levin Y, Savidor A, Kartvelishvily E, Minsky A. 2016. Efficiency in complexity: composition and dynamic nature of mimivirus replication factories. J Virol 90: 10039–10047. https://doi.org/10.1128/JVI.01319-16.
15. Kuznetsov YG, Klose T, Rossmann M, McPherson A. 2013. Morphogenesis of mimivirus and its viral factories: an atomic force microscopy study of infected cells. J Virol 87:11200–11213. https://doi.org/10.1128/JVI.01372-13.
16. Arantes TS, Rodrigues RAL, dos Santos Silva LK, Oliveira GP, de Souza HL, Khalil JYB, de Oliveira DB, Torres AA, da Silva LL, Colson P, Kroon EG, da Fonseca FG, Bonjardim CA, La Scola B, Abrahão JS. 2016. The large marseillevirus explores different entry pathways by forming giant infectious vesicles. J Virol 90:5246–5255. https://doi.org/10.1128/JVI.00177-16.
17. Yuan N, Song L, Zhang S, Lin W, Cao Y, Xu F, Fang Y, Wang Z, Zhang H, Li X, Wang Z, Cai J, Wang J, Zhang Y, Mao X, Zhao W, Hu S, Chen S, Wang J. 2015. Bafilomycin A1 targets both autophagy and apoptosis pathways in pediatric B-cell acute lymphoblastic leukemia. Haematologicaae 100: 345–356. https://doi.org/10.3324/haematol.2014.113324.
18. Mutsafi Y, Shimoni E, Shimon A, Minsky A. 2013. Membrane assembly during the infection cycle of the giant mimivirus. PLoS Pathog 9:e1003367. https://doi.org/10.1371/journal.ppat.1003367.
19. Cerritelli ME, Trus BL, Smith CS, Cheng N, Conway JF, Steven AC. 2003. A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. J Mol Biol 327:1–6. https://doi.org/10.1016/S0022-2836(03)00117-7.
20. La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, Raoult D. 2008. The virophage as a unique parasite of the giant mimivirus. Nature 455: 100–104. https://doi.org/10.1038/nature07218.
21. Desnues C, Raoult D. 2012. Virophages question the existence of satellites. Nat Rev Microbiol 10:234. https://doi.org/10.1038/nrmicro2676-c3.
22. Gaia M, Benamar S, Boughalmi M, Pagnier I, Croce O, Colson P, Raoult D, La Scola B. 2014. Zamilon, a novel virophage with Mimiviridae host specificity. PLoS One 9:e94923. https://doi.org/10.1371/journal.pone.0094923.
23. Gong C, Zhang W, Zhou X, Wang H, Sun G, Xiao J, Pan Y, Yan S, Wang Y. 2016. Novel virophages discovered in a freshwater lake in China. Front Microbiol 7:5. https://doi.org/10.3389/fmicb.2016.00005.
24. Abrahão JS, Boratto P, Dornas FP, Silva LC, Campos RK, Almeida GM, Kroon EG, La Scola B. 2014. Growing a giant: evaluation of the virological parameters for mimivirus production. J Virol Methods 207:6–11. https://doi.org/10.1016/j.jviromet.2014.06.001.
25. Colson P, De Lamballerie X, Fournous G, Raoult D. 2012. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 55:321–332. https://doi.org/10.1159/000336562.
26. Netherton CL, Wileman T. 2011. Virus factories, double membrane vesicles and viroplasm generated in animal cells. Curr Opin Virol 1:381–387. https://doi.org/10.1016/j.coviro.2011.09.008.
27. Siddiqui R, Khan N. 2012. Biology and pathogenesis of Acanthamoeba. Parasit Vectors 5:6. https://doi.org/10.1186/1756-3305-5-6.
28. Suttle CA. 2007. Marine viruses: major players in the global ecosystem. Nat Rev Microbiol 5:801–812. https://doi.org/10.1038/nrmicro1750.
29. Maruri-Avidal L, Domi A, Weisberg AS, Moss B. 2011. Participation of vaccinia virus L2 protein in the formation of crescent membranes and immature virions. J Virol 85:2504–2511. https://doi.org/10.1128/JVI.02505-10.
30. Reed LJ, Muench H. 1938. A simple method of estimating fifty per cent endpoints. Am J Hyg 27:493–497.
31. Khalil JYB, Robert S, Reteno DG, Andreani J, Raoult D, La Scola B. 2016. High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. Front Microbiol 7:26. https://doi.org/10.3389/fmicb.2016.00026.

4.3. Artigo #3: New isolates of pandoraviruses: contribution to the study of replication cycle steps

Os pandoravírus são vírus de dimensões micrométricas e possuem os genomas mais extensos já descritos na virosfera. Muitos trabalhos desde a sua primeira descrição em 2013 têm focado em desvendar aspectos evolutivos a respeito destes vírus. Entretanto, no que se refere aos aspectos biológicos e sobre o seu ciclo de multiplicação existem poucas informações disponíveis. O presente estudo pretende compreender melhor algumas etapas do ciclo de multiplicação destes vírus como a penetração, a morfogênese e a liberação. Para isso, foram realizados ensaios biológicos utilizando três novos isolados de pandoravírus obtidos a partir de amostras ambientais brasileiras. Os resultados mostraram que a infecção causada pelos pandoravírus na ameba *A. castellanii* gera profundas modificações na organização celular com a formação de uma grande fábrica viral elétron-luscente ocupando aproximadamente um terço do citoplasma amebiano. Por meio de análises de imagens de MET foi observado o recrutamento de estruturas de aparência membranosa e de mitocôndrias para o interior e ao redor das fábricas virais, além da degradação nuclear ao longo da infecção. Durante a morfogênese viral foi possível notar que a montagem das partículas pode ser iniciada por ambas as extremidades e não apenas pela região do ostíolo apical, como sugerido em trabalhos anteriores. Por meio da contagem das partículas virais no sobrenadante da cultura durante o curso da infecção, foi observado que estas são liberadas antes da lise celular, indicando que o processo de exocitose também contribui para liberação viral, juntamente com a lise celular. Outros dados que reforçam esta ideia é a observação de partículas recém formadas sendo envoltas por estruturas semelhantes a exossomos em imagens de MET e o tratamento das células infectadas com brefeldina mostrou que o tráfego de membranas afeta a liberação de dois dos três isolados analisados. Em conjunto, os dados obtidos neste estudo fornecem novas informações a respeito do ciclo de multiplicação dos pandoravírus.

Este artigo foi publicado no periódico ***Journal of Virology*** em fevereiro de 2019

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

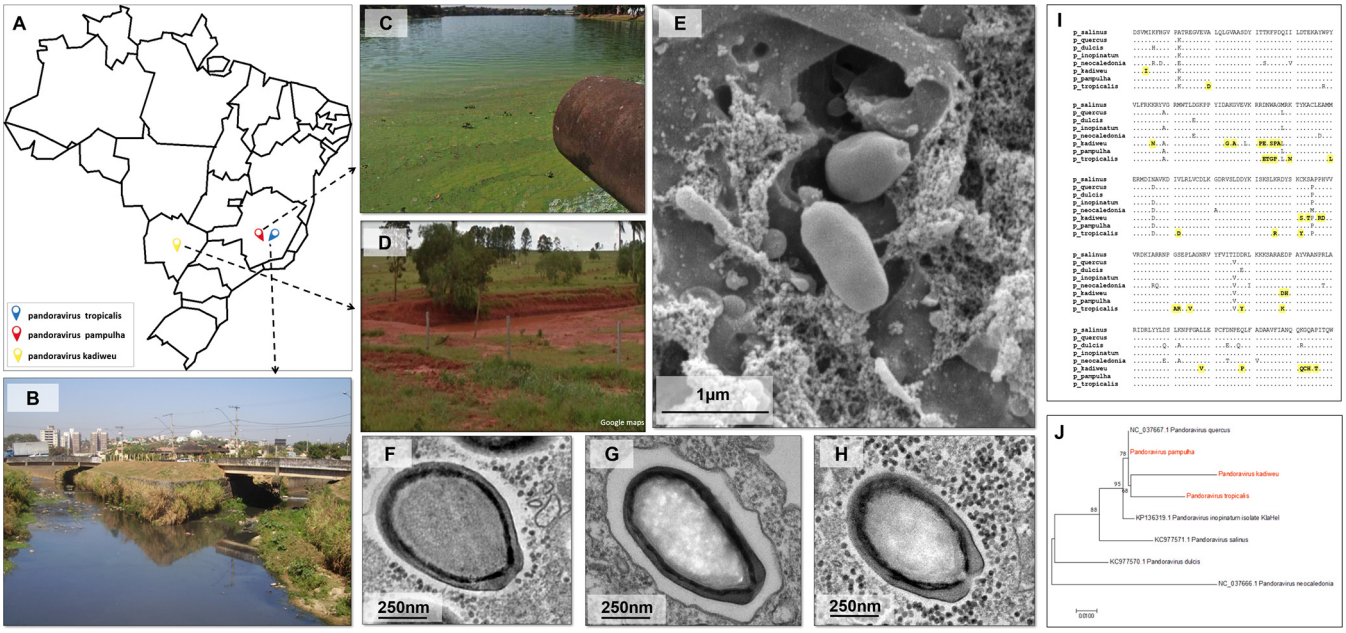


FIG 1 Sites of collection and electron microscopy and phylogenetic analysis of the pandoravirus isolated in this work. (A) Map of Brazil showing where the samples were collected for the isolation of pandoraviruses. (B to D) Representative pictures from the areas of collection: Bom Jesus Creek (B), Mergulhão 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

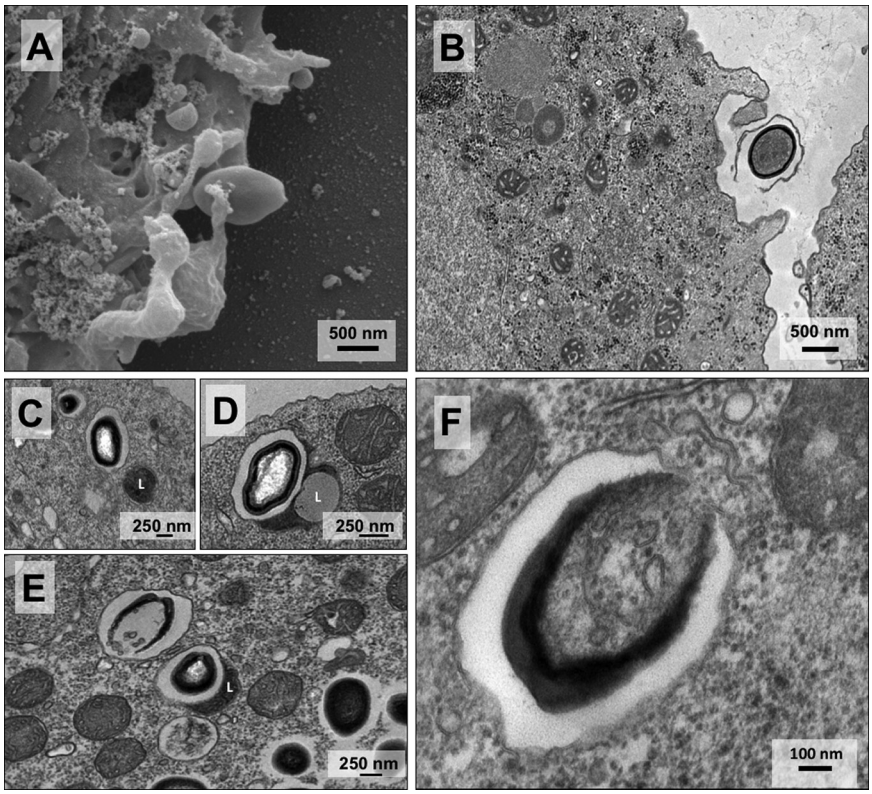


FIG 2 Initial steps of the pandoravirus replication cycle inside the amoebal host. (A and B) Scanning electron microscopy (A) and transmission electron microscopy (B) images show pandoravirus particles entering *Acanthamoeba castellanii* cells, likely as a consequence of phagocytosis. (C) The amoebas 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

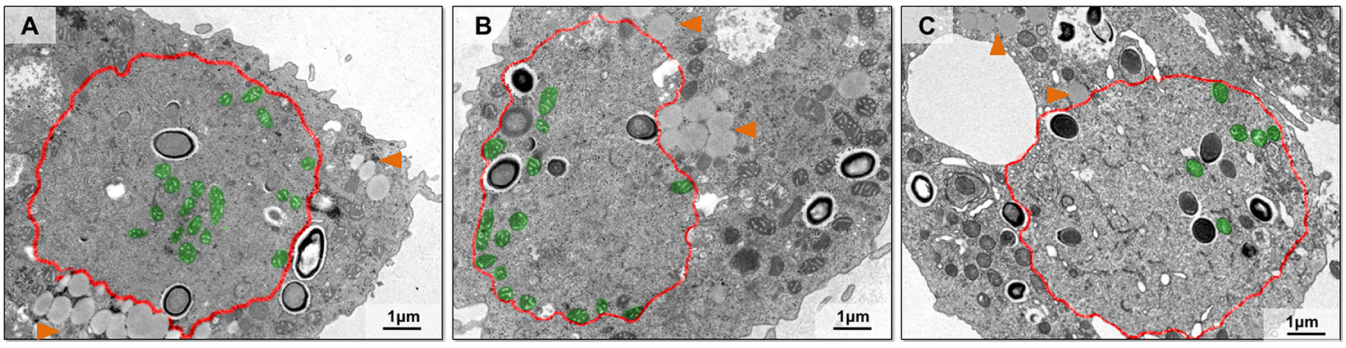


FIG 3 Characterization of pandoravirus viral factories. Viral factories of (A) *Pandoravirus tropicalis*, (B) *P. pampulha*, and (C) *P. kadiweu* were observed by transmission electron microscopy. The region of the viral factories is highlighted in red, the mitochondria present in the interior of the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

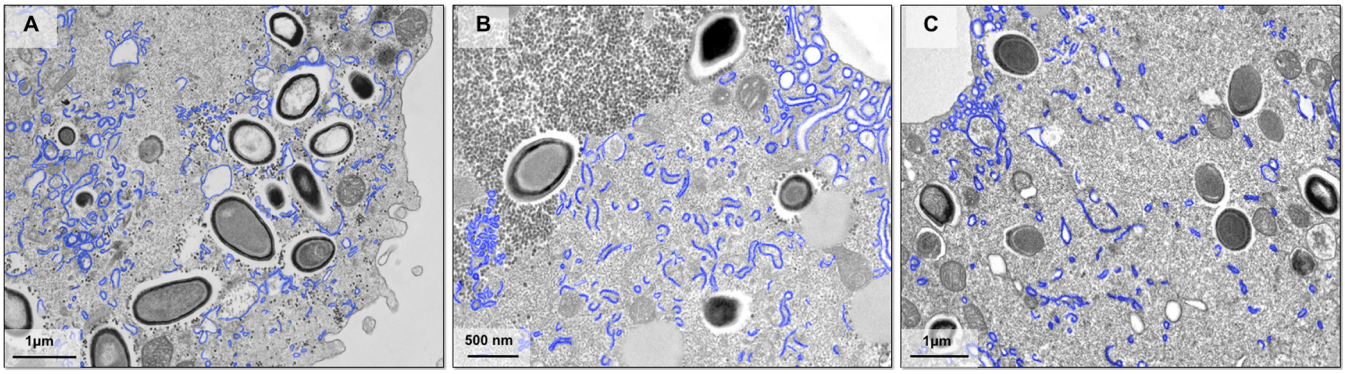


FIG 4 Membranes recruited inside pandoravirus viral factories. (A) Transmission electron microscopy of *P. tropicalis* viral factories. (B) Transmission electron microscopy of *P. pampulha* viral factories. (C) Transmission electron microscopy of *P. kadiweu* viral factories. The membranes recruited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

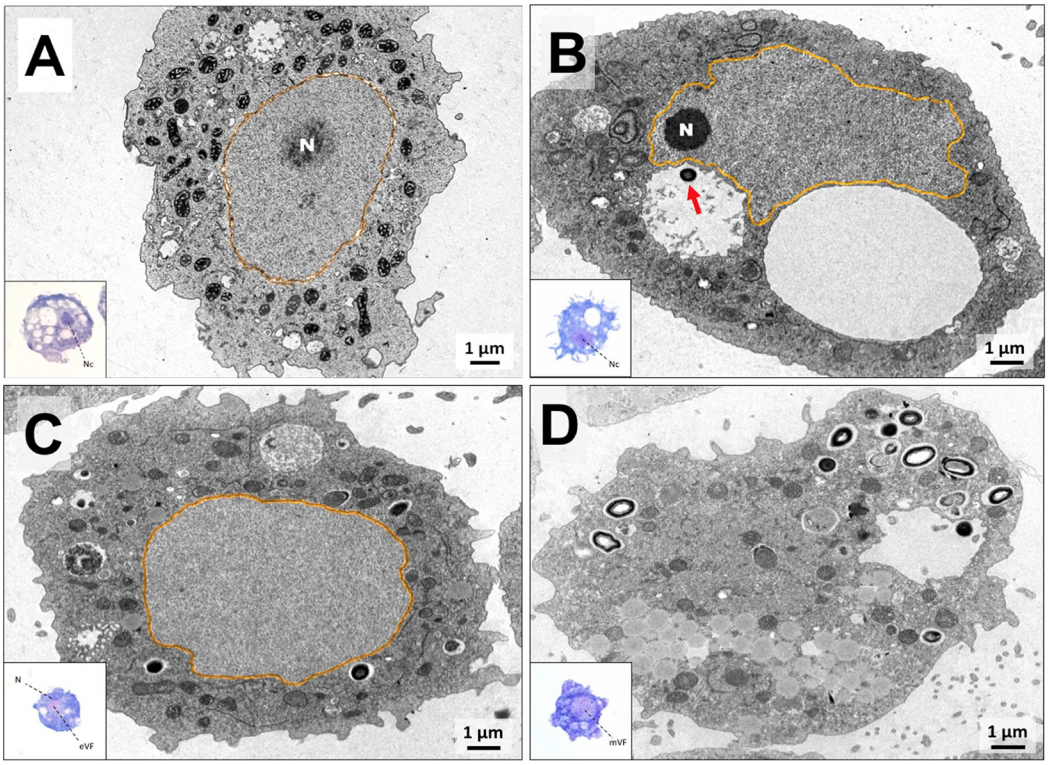


FIG 5 The *Acanthamoeba castellanii* cell nucleus becomes disorganized and loses its natural shape during the course of pandoravirus infection. (A) Transmission electron microscopy image showing a noninfected *Acanthamoeba castellanii* cell and how its nucleus is normally organized in this situation; 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.

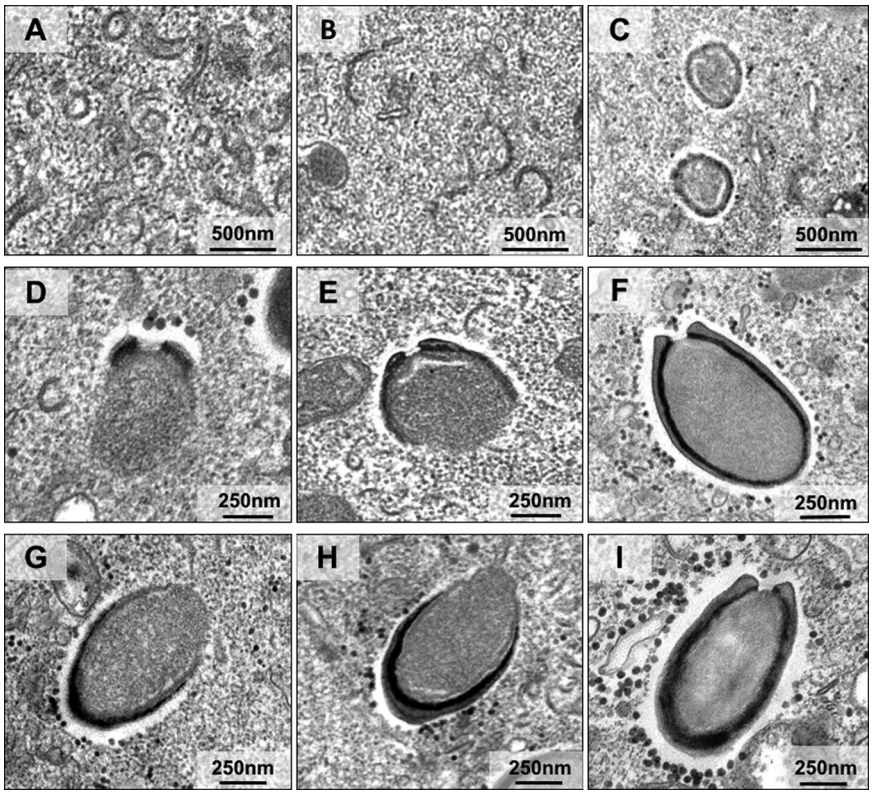


FIG 6 Morphogenesis of pandoravirus particles. Transmission electron microscopy images show stages of pandoravirus particle formation. (A to C) Crescent-like structures with different sizes, inside the viral factory, growing in thickness and complexity. (D to F) Particles being formed from the ostiolo-like 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-

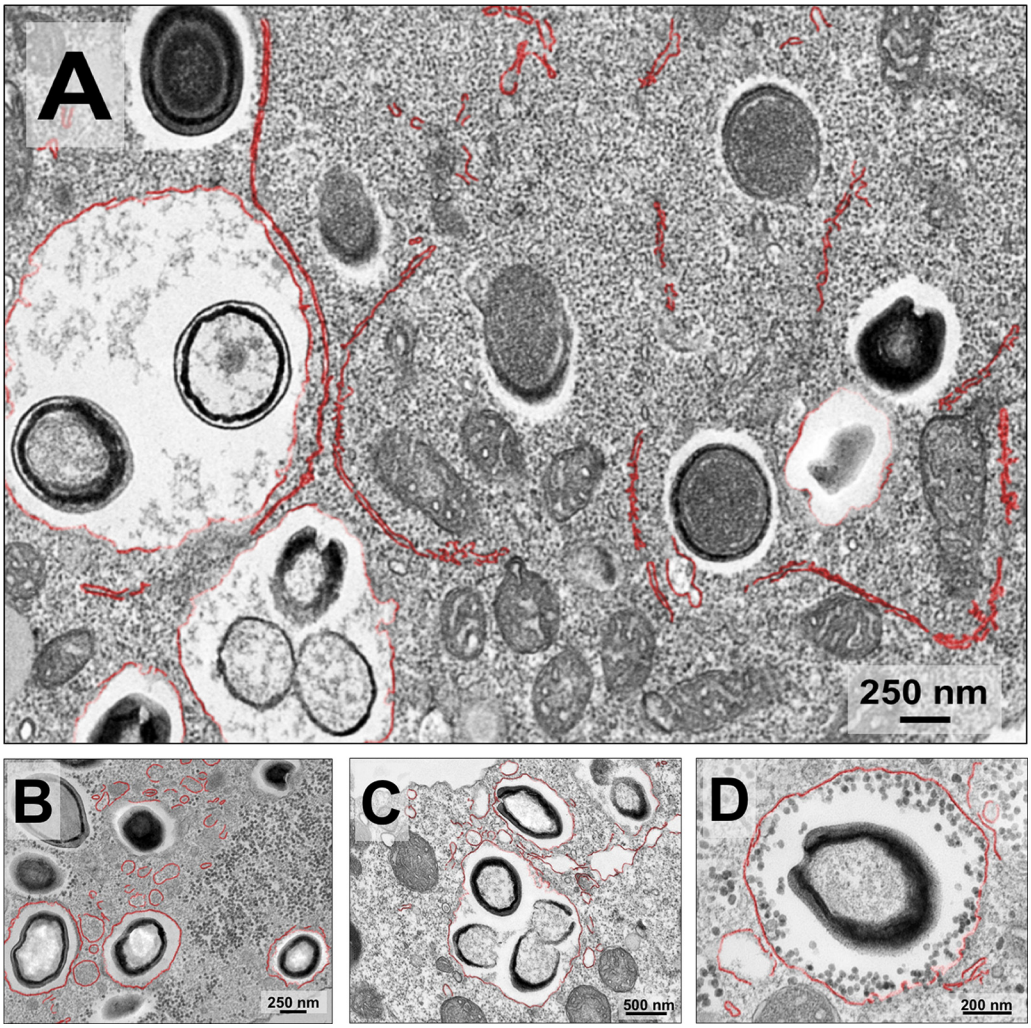


FIG 7 Transmission electron microscopy images showing pandoravirus particles being packaged into exosomes. (A) The late steps in pandoravirus replication are marked by intense membrane trafficking in the cytoplasm of the amoebal host (highlighted in red). This event is easily observed around the viral 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.

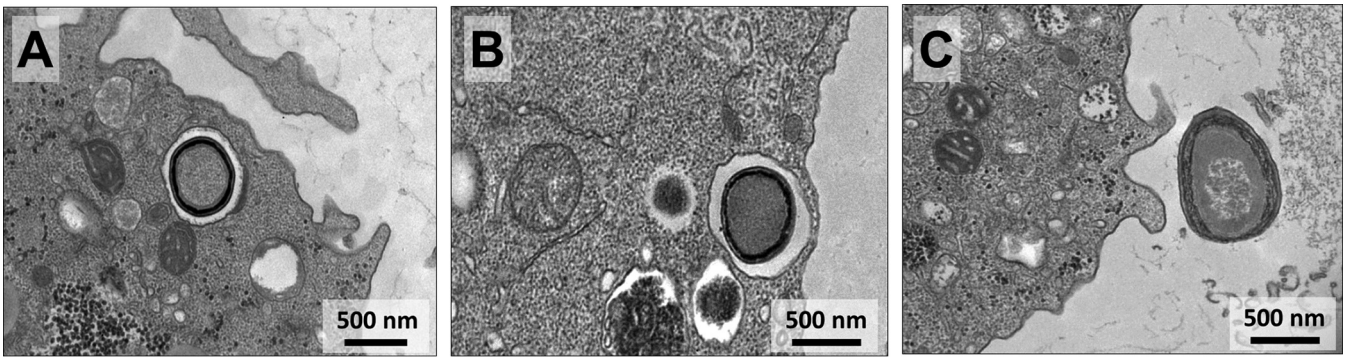


FIG 8 Transmission electron microscopy images demonstrating sequential steps of pandoravirus particle exocytosis. The images demonstrate that in late stages of infection the particles of pandoravirus start being packaged inside vesicles (A and B), becoming closer to the cytoplasmic membrane of the host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

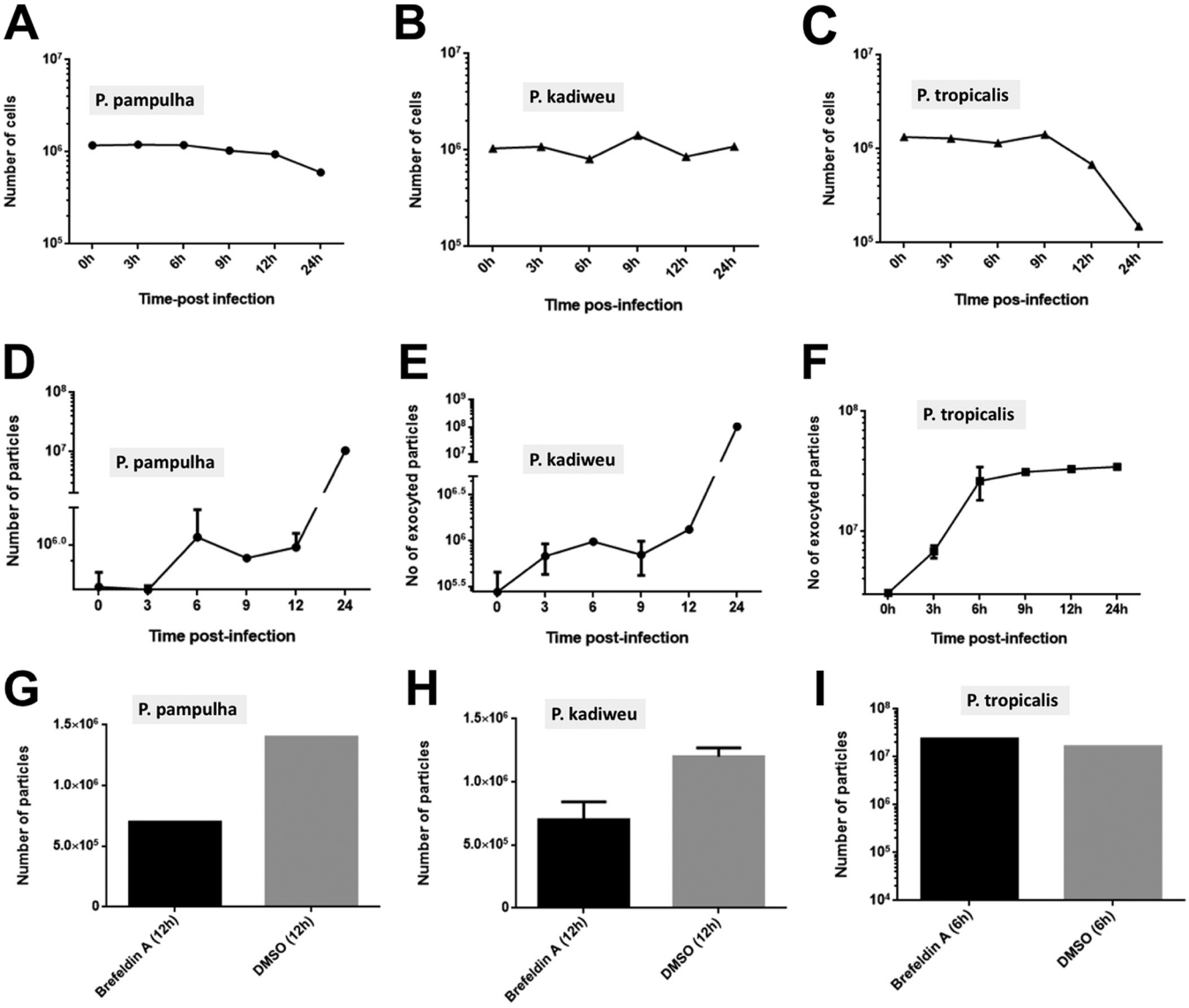


FIG 9 Analysis of the time course of infection for the pandoravirus isolates in *Acanthamoeba castellanii* cultures. The course of infection of *P. tropicalis*, *P. pampulha*, and *P. kadiweu* was established. (A to C) First, we observed the kinetics for the diminishing of the number of amoebas during 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.

5. DISCUSSÃO

As amebas de vida livre são conhecidas como organismos ubíquos no planeta (TANVEER *et al.*, 2015; SHEID, 2018). Como elas são hospedeiras dos vírus gigantes, foi sugerido que estes vírus também estivessem amplamente distribuídos pelo mundo (AHERFI *et al* 2016). Para responder sobre esta e outras questões a respeito da diversidade e distribuição dos vírus gigantes, os estudos de metagenômica e de prospecção são fundamentais. Neste trabalho, apresentamos o isolamento de 68 vírus gigantes utilizando amostras coletadas no Brasil e na Antártica. Os resultados obtidos neste estudo apontaram o maior sucesso de isolamento (91,17% dos isolados) de mimivírus em relação aos demais grupos de vírus gigantes, corroborando estudos anteriores que também apontaram o maior isolamento deste grupo viral (LA SCOLA *et al*., 2010; ANDRADE *et al*., 2014; DORNAS *et al*., 2015). Além disso, o insucesso no isolamento viral a partir de amostras clínicas, também condiz com o número reduzido de isolados clínicos comparado aos isolados ambientais de vírus descritos na literatura (revisado por COLSON *et al*., 2016; 2017; DARE et al., 2008; SAADI *et al*., 2013).

Este trabalho também descreve o isolamento de dois novos isolados de marseillevírus e do primeiro cedratvírus brasileiro. Uma análise aprofundada do ciclo de multiplicação do cedratvírus isolado neste trabalho foi realizada e forneceu informações importantes a respeito do ciclo de multiplicação destes vírus (SILVA *et al.,* 2018). Até o momento apenas o brazilian marseillevírus e o golden marseillevírus já haviam sido descritos no Brasil. Estes dois vírus apresentam uma alta diversidade genômica, e foi proposto que eles representem duas linhagens diferentes dentro da família *Marseilleviridae* (SANTOS *et al.,* 2016; DORNAS *et al.*, 2016). Sendo assim, análises posteriores a respeito do genoma e a biologia dos novos isolados de marseillevírus apresentados neste estudo também podem contribuir no entendimento da diversidade desta família viral.

Em relação aos pandoravírus, desde a sua primeira descrição em 2013, estudos de metagenômica indicaram a presença de marcadores genéticos associados a estes vírus em diferentes ambientes como carcaça de animais, plasma humano, água, solo e mosquitos (PHILIPPE *et al*., 2013; VERNEAU *et*

al., 2016; TEMMAN *et al*., 2015; 2017; ATONI *et al*., 2018; BRINKMAN *et al.,* 2018). Entretanto, apenas doze isolados de pandoravírus já foram descritos em estudos utilizando amostras de diferentes partes do mundo, incluindo as amostras descritas no presente trabalho (DORNAS *et al*., 2015; LEGENDRE et al 2018; AHERFI et al 2018). Neste contexto é importante destacar a importância dos trabalhos de prospecção brasileiros, já que seis isolados entre os doze pandoravírus já descritos são brasileiros e três foram descritos no presente trabalho. O isolamento dessas novas amostras de pandoravírus contribuíram para realização de trabalhos que analisaram o pangenoma, as relações filogenéticas e evolutivas destes vírus (AHERFI *et al*., 2018).

Estudos anteriores de metagenômica e prospecção viral também relataram e/ou indicaram a presença dos vírus gigantes em todos os continentes incluindo em regiões de condições ambientais extremas como em amostras da Antártica, de desertos e de lagoas salinas (ABRAHÃO *et al*., 2018; KEREPESI, GROLMUSZ, 2014; 2017). Amebas já foram isoladas e sequências genômicas relacionadas à família *Mimiviridae* e a virófagos também já foram descritas em amostras da Antártica o que é um indicativo da presença de vírus gigantes neste continente (BROWN *et al*., 1982; YAU *et al.,* 2011; ROYLES *et al.,* 2013; KEREPESI, GROLMUSZ, 2017). No entanto, até onde sabemos, este trabalho relata o primeiro isolamento de vírus gigantes associado a amebas de amostras da Antártica, confirmando algumas expectativas anteriores da ubiquidade destes vírus.

A técnica escolhida para obtenção dos isolados foi a inoculação direta, por ser uma técnica não fastidiosa e já ter sido utilizada com sucesso em estudos anteriores de prospecção de vírus gigantes (DORNAS *et al*., 2015). Para a identificação dos isolados foram utilizadas diferentes técnicas e foi demonstrado que a utilização apenas da PCR muitas vezes não foi suficiente para identificação, o que pode ser explicado pela alta variabilidade genética entre esses vírus (ASSIS *et al.,* 2015; 2017; BORATTO *et al*., 2018; LEGENDRE *et al.,* 2018; AHERFI *et al.,* 2018). Também é importante considerar que entre os isolados não classificados em nenhuma das três linhagens conhecidas podem haver representantes de novas linhagens que ainda não foram descritas. Os resultados obtidos na prospecção demonstram a

importância de se utilizar um conjunto de técnicas (e não uma técnica apenas) para a identificação, conforme realizado neste estudo.

Nos últimos anos, a maior parte dos trabalhos sobre vírus gigantes isolados em diferentes lugares do mundo têm focado na caracterização genômica e evolutiva destes isolados, sendo a caracterização biológica parcialmente explorada (BOYER *et al*., 2009; KOONIN *et al*., 2017; LEGENDRE *et al.,* 2018; AHERFI *et al.,* 2018). Diante dessa necessidade de aprofundamento das análises de como estes vírus se multiplicam e como é sua relação com a célula hospedeira, escolhemos dois grupos virais isolados neste trabalho para avançarmos neste aspecto. O primeiro grupo viral escolhido foi o dos mimivírus. Buscamos investigar alguns aspectos referentes as etapas de penetração, desnudamento e morfogênese dos mimivírus. O tamanho das partículas dos mimivírus (aproximadamente 700 nm) associados ao estilo de vida das amebas de vida livre que se alimentam por fagocitose, juntamente com as imagens de microscopia eletrônica, levaram à suposição de que esses vírus penetram nas células hospedeiras por fagocitose. Entretanto, até o presente trabalho não haviam dados biológicos para validar essas suposições. Os dados obtidos demonstraram que a inibição da fagocitose por citocalasina D, um inibidor previamente utilizado em amebas da espécie *A. castellanii*, capaz de inibir a polimerização de actina, resulta em uma redução significativa da incorporação de partículas de mimivírus, o que corrobora a hipótese de que o principal mecanismo de penetração utilizado por esse vírus é a fagocitose (ALSAM *et al*., 2005; CHRISMAN; ALVAREZ; CASADEVALL, 2010; SOTO-ARREDONDO *et al*., 2014; TAYLOR *et al*., 1995). Esta inibição não foi observada durante a infecção pelos marseillevírus, já que foi demonstrado que estes vírus são capazes de utilizar vias outras vias endocíticas para penetração (ARANTES *et al.,* 2016).

Após a penetração viral, o desnudamento ocorre após a abertura de um vértice do capsídeo denominado *stargate* e a formação de um portal pela fusão da membrana interna viral com a membrana do fagossomo, permitindo que a semente viral seja liberada no citoplasma (ZAUBERMAN *et al.,* 2008). Neste trabalho, foi investigado se a acidificação do endossomo é um fator importante no processo de desnudamento. A utilização da bafilomicina (inibidor da acidificação de endossomo) resultou em diminuição da replicação viral. Além

disso, as análises das imagens de MET mostram a ocorrência da fusão de uma estrutura semelhante a lisossomo ao fagossomo. É possível que uma diminuição no pH após a formação do fagolisossomo seja o fator desencadeante da abertura do *stargate* dos mimivírus, mas permanece incerto qual pH é ideal e se as enzimas presentes nos lisossomos são essenciais para este processo.

Os processos que ocorrem após a liberação da semente viral no citoplasma celular até o início da morfogênese são pouco conhecidos e algumas vezes controversos. É possível que fatores nucleares estejam envolvidos no início da replicação do DNA viral, mas ainda não há evidências suficientes que comprovem esta hipótese (SUZAN-MONT *et al.,* 2007; MUTSAFI *et al.,* 2010). A morfogênese viral ocorre dentro de grandes FV elétron-densas formadas no citoplasma celular. Estudos anteriores mostraram que as membranas do retículo endoplasmático (RE) teriam papel de suporte na montagem do capsídeo em formação, o qual ocorreria pela aquisição de unidades pentaméricas individuais em sequência (MUTSAFI *et al*., 2013). Nossos dados corroboram essa teoria, demonstrando a existência e a expansão de estruturas lamelares, de forma análoga à dos crescentes descritos para poxvírus e marseillevírus (ARANTES *et al*., 2016; MARURI-AVIDAL *et al*., 2011).

Essas estruturas lamelares crescentes aumentam de tamanho pela incorporação de proteínas que se afastam da região mais elétron-densa da FV, até passarem por uma área de aspecto fibrilar, que denominamos área de aquisição de fibrila (AAF). Essa região recebeu este nome porque foi proposto que é nesta área que as novas partículas virais adquirem fibrilas superficiais. Imagens de MET mostram que partículas em formação que estão na interface da AAF com o citoplasma celular apresenta fibrilas na porção dos capsídeos que já passaram pela AAF, enquanto que a porção do capsídeo que ainda permanece imersa na AAF não possui fibrilas. Além disso, a AAF parece possuir maior dimensão nas FV iniciais (em células apresentando poucas partículas formadas) e está reduzida nas células que já apresentam muitas partículas formadas, sugerindo que essa matriz fibrilar é consumida continuamente à medida que as partículas passam por ela. Curiosamente, as proteínas fibrilares não foram detectadas na FV isolada sem a região AAF

analisada por meio de abordagens proteômicas (FRIDMANN-SIRKIS *et al*., 2016). As análises iniciais da AAF nas amostras M4 e Borely moumouvirus mostraram que ela está presente na fábrica viral destes isolados como era esperado, uma vez que estes vírus também apresentam fibrilas, apesar de elas estarem presentes em menor quantidade (Figura 2B e C). Não temos uma amostragem de imagens grandes o suficiente para determinar diferenças no tamanho das AAF formadas para cada amostra viral, e também seria importante a aferição do tamanho das AAF após diferentes tempos de infecção para uma análise mais detalhada. Futuros estudos poderão incluir a caracterização das AAF no estudo do ciclo de multiplicação de diferentes amostras e linhagens de mimivírus.

As análises deste trabalho demonstram que a aquisição de fibrilas ocorre na periferia das FV dos mimivírus e não na periferia celular como proposto no modelo anterior, e que a aquisição de fibrilas e de genoma pode ocorrer simultaneamente, e não necessariamente de forma sequencial como proposto anteriormente (KUZNETSOV *et al*., 2013). Apesar de já ter sido demonstrado para maior parte dos vírus conhecidos, que durante a multiplicação viral uma parte da progênie formada é composta por partículas defectivas, no caso dos vírus gigantes a presença de partículas defectivas estava usualmente associada a presença de virófagos (LA SCOLA *et al*., 2008; DESNUES *et al*., 2012). Neste trabalho foi demonstrado que mesmo na ausência de virófagos, partículas defectivas também compõe parte da progênie viral gerada, mesmo quando se é utilizada uma MOI baixa, como demonstrado também por Abrahão e colaboradores em 2014. As análises apresentadas neste estudo, juntamente com dados publicados anteriormente, resultaram na proposição de um modelo mais atualizado a respeito do ciclo de multiplicação dos mimivírus.

Após a caracterização de algumas etapas do ciclo de multiplicação dos mimivírus, a presente tese também teve como objetivo analisar algumas etapas do ciclo de multiplicação dos pandoravírus, fornecendo informações relevantes para o melhor entendimento da biologia de três isolados brasileiros, denominados P. pampulha, P. tropicalis, e P. kadiweu. A análise filogenética realizada utilizando uma região do gene da DNA polimerase B destes vírus mostrou que estes vírus apresentam diferenças genéticas entre si. Porém

futuros estudos determinarão se eles pertencem a diferentes clados, já que até o momento, não existem regras ou parâmetros oficiais disponíveis para estabelecimento de separação de grupos dentro da hipotética família Pandoraviridae.

Assim como foi hipotetizado para os mimivírus, também foi sugerido que os pandoravírus penetram nas amebas por fagocitose (PHILIPPE *et al*., 2013). As análises das imagens de microscopia apresentadas neste trabalho reforçam essa hipótese, pois partículas podem ser vistas dentro de grandes vesículas no citoplasma da ameba dentro de 30 min pós-infecção. Imagens de microscopia eletrônica de varredura de tempos iniciais de infecção também mostram a formação de pseudópodes englobando partículas de pandoravírus. Apesar desta evidência, a possibilidade de utilização do processo de macropinocitose para penetração viral não pode ser anulada apenas pelos dados de microscopia que foram obtidos (MERCER *et al*., 2008). A via de macropinocitose foi investigada como possível via de penetração dos cedratvirus utilizando o inibidor desta via, 5-(*N*-ethyl-*N*-isopropyl) amiloride (EIPA). Este inibidor não afetou a penetração de partículas de cedratvírus, porém o seu efeito em células de *Acanthamoeba* ainda não foi profundamente estudado e a utilização da macropinocitose na penetração de cedratvírus também não pode ser descartada (SILVA *et al*., 2018).

O desnudamento dos pandoravírus foi descrito em trabalhos anteriores como sendo semelhante ao processo descrito para os mimivírus. Também é observada a formação de um canal na região do ostíolo apical pela fusão da membrana interna da partícula com a membrana do endossomo (PHILIPPE *et al*., 2013; LEGENDRE *et al*., 2018). As FVs dos pandoravírus foram caracterizadas neste trabalho e foi observado um intenso recrutamento de mitocôndrias para o interior e ao redor das FVs, fenômeno que havia sido descrito anteriormente para os cedratvirus (SILVA *et al*., 2018). A presença de mitocôndrias nessa região poderia estar relacionada a otimização da aquisição de energia durante da replicação viral (NOVOA *et al*., 2005). Também faz parte dessa reorganização celular ao longo da infecção, a degradação nucleolar e nuclear, como já observado para outros pandoravírus (LEGENDRE *et al*., 2018). Entretanto, descrevemos esta etapa de forma mais detalhada mostrando que primeiro ocorre a degradação nucleolar no tempo entre 3 e 6

h.p.i e nos tempos mais tardios de infecção, a partir de 9 h.p.i, o núcleo desaparece completamente. Análises do conteúdo gênico de P. salinus realizada por Philippe e colaboradores em 2013 mostraram que este vírus não apresenta genes que codificam componentes essenciais na replicação do DNA, como DNA ligases e topoisomerases. Foi sugerido então que a multiplicação dos pandoravírus requer componentes celulares normalmente presentes no núcleo da ameba (PHILIPPE *et al.,* 2013).

A morfogênese dos pandoravírus parece ser iniciada a partir de estruturas semicirculares elétron-densas, semelhantes às crescentes observadas na montagem de outros vírus do grupo dos NCLDV, como o vaccinia vírus, mimivírus, marseillevírus e cedratvírus (ARANTES *et al*., 2016; SUAREZ *et al*., 2013; SILVA *et al*., 2018). Nas etapas subsequentes da montagem das partículas virais pode ser observado que o capsídeo e o conteúdo interno das partículas são montados simultaneamente, semelhantes ao que foi descrito para os mollivirus, pithovirus e os outros pandoravírus (PHILIPPE *et al.,* 2013; LEGENDRE *et al.*, 2014; LEGENDRE *et al*., 2015). Entretanto, diferente do foi previamente sugerido, observamos que as partículas virais parecem ser montadas a partir de ambas as extremidades e não apenas pelo ostíolo apical (PHILIPPE *et al.,* 2013). As semelhanças observadas com relação à organização de FV e a morfogênese do pandoravírus e outros vírus do grupo NCLDV reforçam a ideia sugerida anteriormente de que esses vírus compartilham um ancestral comum (SHARMA *et al.,* 2015; YUTIN *et al.,* 2013).

Em um estudo publicado recentemente, diferentes isolados de pandoravírus apresentaram diferentes tempos de liberação viral. Além disso, foi sugerido neste estudo o que a exocitose também é utilizada como via na liberação viral (LEGENDRE *et al*., 2018). Os padoravírus brasileiros analisados também apresentaram variações no tempo de liberação viral durante a infecção e diferentes respostas diante do tratamento com brefeldina. Assim como observado por Legendre e colaboradores em 2018, os dados gerados neste estudo também sugerem a presença de diversidade genética e biológica entre os isolados. A análise da etapa de liberação viral dos pandoravírus brasileiros corroboram a hipótese de que uma parte da progênie é exocitada já que partículas de pandoravírus foram observadas no sobrenadante da cultura

infectada em tempos anteriores à lise celular. Além disso, uma série de imagens de MET também sugerem que as partículas virais recém formadas são envolvidas por estruturas membranosas semelhantes a exossomos. O fato da brefeldina afetar a liberação viral de dois dos três pandoravírus analisados também indica o papel da exocitose uma vez que este processo é dependente do tráfego de membranas (SILVA *et al.,* 2018).

O isolamento e a caracterização de novos vírus gigantes são objetivos centrais do nosso grupo de pesquisa. Em conjunto, os resultados obtidos nesta tese contribuem com novas informações a respeito da distribuição, diversidade e no entendimento de alguns aspectos do ciclo de replicação dos mimivírus e pandoravírus. O isolamento de novas espécies de vírus gigantes e futuros estudos de caracterização biológica, especialmente a nível molecular, poderão avançar ainda mais na compreensão da biologia desses vírus complexos.

6. CONCLUSÕES

- Este estudo relata o isolamento:
  - Do primeiro mimivírus a partir de amostras da Antártica;
  - De três pandoravírus dos 12 que já foram isolados no mundo;
  - Do terceiro cedratvírus isolado do mundo e o primeiro do Brasil;
  - Do terceiro marseillevírus isolado do Brasil.

- A maioria dos novos isolados foram identificados como mimivírus;

- O maior número total de isolados foi obtido a partir das amostras de água, porém a partir das amostras de esgoto foi possível obter maior número de grupos diferentes de vírus gigantes;

- A associação de técnicas como, PCR, microscopia óptica e eletrônica, foi essencial para identificação de todos os isolados obtidos;

- Os resultados apresentados neste trabalho reforçam a ideia de que os vírus gigantes são ubíquos assim como seus hospedeiros protistas;

- Ensaios biológicos com inibidores associados com a microscopia eletrônica sugerem que a penetração dos mimivírus na célula hospedeira ocorre por meio de fagocitose;

- O capsídeo dos mimivírus parece ser montado a partir de estruturas lamelares crescentes observadas nas imagens de microscopia eletrônica;

- Na periferia das fábricas virais, formadas durante o ciclo de multiplicação dos mimivírus foi identificada uma região que fornece fibrilas às partículas que estão sendo montadas, sendo esta região denominada de área de aquisição de fibrilas;

- Foi observado que aquisição de genoma e de fibrilas pelas partículas de mimivírus pode ocorrer simultaneamente e não apenas de forma sequencial, como relatado anteriormente;

- A formação de partículas defectivas durante o ciclo de multiplicação dos mimivírus ocorre mesmo na ausência de virófagos;

- A fábrica viral formada durante a multiplicação dos pandoravírus é caracterizada como largas regiões elétron-luscentes e ocupa cerca de 1/3 do citoplasma amebiano;

- Foi observado que mitocôndrias são recrutadas ao redor e no interior das fábricas virais dos pandoravírus;

- Foi demonstrado um intenso recrutamento de estruturas de aparência membranosas no interior das fábricas virais dos pandoravírus. Sugerimos que este recrutamento é importante na morfogênese viral e na formação dos exossomos;

- Foi descrito que durante o ciclo de multiplicação dos pandoravírus a morfologia nuclear sofre alterações e que o desaparecimento nucleolar precede a degradação do restante do núcleo;

- A morfogênese das partículas dos pandoravírus pode ser iniciada pelas duas extremidades da partícula e não apenas pela extremidade que contém o ostíolo apical, como sugerido anteriormente;

- Foi observado a presença de estruturas similares a exossomos envolvendo as partículas virais recém formadas no citoplasma das amebas infectadas com pandoravírus;

- A inibição do tráfego de membranas impactou na liberação viral de dois dos três isolados de pandoravírus analisados, indicando a diversidade biológica dentre os isolados;

- A exocitose é um dos processos de liberação viral utilizada pelos pandoravírus, juntamente com a lise celular.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHÃO, J. S. et al. Acanthamoeba polyphaga mimivirus and other giant viruses:An open field to outstanding discoveries. Virology Journal, v. 11, n. 1, p. 1–12, 2014.

ABRAHÃO, J. S. et al. The analysis of translation-related gene set boosts debates around origin and evolution of mimiviruses. PLoS Genetics. [S.l: s.n.]. , 2017.

ABRAHÃO, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. Nature Communications, v. 9, n. 1, p. 749, 2018.

AHERFI, S.; COLSON, P.; LA SCOLA, B.; et al. Giant viruses of amoebas: An update. Frontiers in Microbiology, v. 7, n. MAR, p. 1–14, 2016.

AHERFI, S., ANDREANI, J., BAPTISTE, E., OUMESSOUM, A., DORNAS, F.P., ANDRADE, A., CHABRIERE, E., ABRAHAO, J., LEVASSEUR, A., RAOULT, D., LA SCOLA, B., COLSON, P. A Large Open Pangenome and a Small Core Genome for Giant Pandoraviruses. Frontiers in microbiology, v. 9, n.1486, p. 1-13, 2018.

ANDRADE, K. R. et al. Oysters as hot spots for mimivirus isolation. Archives of Virology, v. 160, n. 2, p. 477–482, 2014.

ANDREANI, J. et al. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. Viruses, v. 8, n. 11, p. 1–11, 2016.

ANDREANI, J. et al. Pacmanvirus, a new giant icosahedral virus at the crossroads between Asfarviridae and Faustoviruses. Journal of Virology, p. JVI.00212-17, 2017.

ANDREANI, J. et al. Orpheovirus IHUMI-LCC2: A new virus among the giant viruses. Frontiers in Microbiology, v. 8, p. 1–11, 2018

ANTWERPEN, M. H. et al. Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. Genome announcements, v. 3, n. 2, p. e00136-15, 2015.

ARSLAN, D. et al. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proceedings of the National Academy of Sciences, v. 108, n. 42, p. 17486–17491, 2011.

ASSIS, F. L. et al. Pan-genome analysis of Brazilian lineage a amoebal mimiviruses. Viruses, v. 7, n. 7, p. 3483–3499, 2015.

ASSIS, F. L. et al. Genome characterization of the first mimiviruses of lineage C isolated in Brazil. Frontiers in Microbiology, v. 8, n. DEC, p. 1–11, 2017

ATONI E, WANG Y, KARUNGU S, WARUHIU C, ZOHAIB A, OBANDA V, AGWANDA B, MUTUA M, XIA H, YUAN Z. Metagenomic virome analysis of Culex mosquitoes from Kenya and China. Viruses 10, 2018.

BAJRAI, L. H.; BENAMAR, S.; et al. Kaumoebavirus, a new virus that clusters with Faustoviruses and Asfarviridae. Viruses, v. 8, n. 11, 2016.

BENAMAR, S. et al. Faustoviruses: Comparative genomics of new megavirales family members. Frontiers in Microbiology, v. 7, n. FEB, p. 1–9, 2016.

BORATTO, P. V. M. et al. Niemeyer virus: A new mimivirus group A isolate harboring a set of duplicated aminoacyl-tRNA synthetase genes. Frontiers in Microbiology, v. 6, n. NOV, p. 1–11, 2015.

BORATTO PVM, DORNAS FP, DA SILVA LCF, RODRIGUES RAL, OLIVEIRA GP, CORTINES JR, DRUMOND BP, ABRAHÃO JS. Analyses of the Kroon Virus Major Capsid Gene and Its Transcript Highlight a Distinct Pattern of Gene Evolution and Splicing among Mimiviruses. Journal of Virology, 92(2); 2018.

BOUGHALMI, M. et al. First isolation of a marseillevirus in the diptera syrphidae eristalis tenax. Intervirology, v. 56, n. 6, p. 386–394, 2013.

BOYER, M. et al. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proceedings of the National Academy of Sciences, v. 106, n. 51, p. 21848–21853, 2009.

BOYER M, AZZA S, BARRASSI L, KLOSE T, CAMPOCASSO A, PAGNIER I, FOURNOUS G, BORG A, ROBERT C, ZHANG X, DESNUES C, HENRISSAT B, ROSSMANN MG, LA SCOLA B, RAOULT D. Mimivirus shows dramatic genome reduction after intraamoebal culture. Proc Natl Acad Sci U S A 108:10296–10301, 2011.

BRINKMAN NE, VILLEGAS EN, GARLAND JL, KEELY SP. Reducing inherent biases introduced during DNA viral metagenome analyses of municipal wastewater. PLoS One 13:e0195350. https://doi.org/10.1371/journal.pone .0195350, 2018.

BROWN,T J; CURSONS, R T M; KEYS, E A. Amoebae from Antarctic Soil and Water. Applied and environmental microbiology,v. 44; p. 491-493, 1982.

CAMPOS, R. K. et al. Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. Virology Journal, v. 11, n. 1, p. 95, 2014.

CHRISMAN, C. J.; ALVAREZ, M.; CASADEVALL, A. Phagocytosis of cryptococcus neoformans by, and nonlytic exocytosis from, Acanthamoeba castellanii. Applied and Environmental Microbiology, v. 76, n. 18, p. 6056–6062, 2010

COLSON, P.; FANCELLO, L.; et al. Evidence of the megavirome in humans. Journal of Clinical Virology, v. 57, n. 3, p. 191–200, 2013.

COLSON, P.; PAGNIER, I.; et al. “Marseilleviridae”, a new family of giant viruses infecting amoebae. Archives of Virology, v. 158, n. 4, p. 915–920, 2013.

COLSON, P. et al. Mimivirus: leading the way in the discovery of giant viruses of amoebae. Nature Reviews Microbiology, 2017.

COLSON, P. et al. The role of giant viruses of amoebas in humans. Current Opinion in Microbiology, v. 31, p. 199–208, 2016.

COLSON, P.; LA SCOLA, B.; RAOULT, D. Giant viruses of amoebae: A journey through innovative research and paradigm changes. Annual Review of Virology, v. 4, n. 1, p. annurev-virology-101416-041816, 2017.

CLOUTHIER S., ANDERSON B, KURATHC G., BREYTAC R. Molecular systematics of sturgeon nucleocytoplasmic large DNA viruses. Molecular Phylogenetics and Evolution, v.128, p. 26–37, 2018.

DARE, R. K.; CHITTAGANPITCH, M.; ERDMAN, D. D. Screening Pneumonia Patients for Mimivirus1. Emerging infectious diseases, v. 14, n. 3, p. 465–467, 2008.

DEEG C.M., CHOW, C. T., SUTTLE, C. A. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. eLife, v. 7, p.1-22, 2018.

DESNUES C, RAOULT D. Virophages question the existence of satellites. Nat Rev Microbiol 10:234, 2012.

DORNAS, F. P. et al. A Brazilian marseillevirus is the founding member of a lineage in family marseilleviridae. Viruses, v. 8, n. 3, 2016.

DORNAS, F. P. et al. Isolation of new Brazilian giant viruses from environmental samples using a panel of protozoa. Frontiers in Microbiology, v. 6, n. OCT, p. 1–9, 2015.

DORNAS, F. P. et al. Mimivirus circulation among wild and domestic mammals, Amazon Region, Brazil. Emerging Infectious Diseases, v. 20, n. 3, p. 469–472,

2014.

DOS SANTOS, R. N. et al. A new marseillevirus isolated in Southern Brazil from Limnoperna fortunei. Scientific Reports, v. 6, n. 1, p. 35237, 2016.

DOS SANTOS SILVA, L.K.; ARANTES, T.S.; ANDRADE, K.R.; LIMA RODRIGUES, R.A.; MIRANDA BORATTO, P.V, DE FREITAS ALMEIDA, G.M.; KROON, E.G.; LA SCOLA, B.; CLEMENTE, W.T.; SANTOS ABRAHÃO, J. (2015). High positivity of mimivirus in inanimate surfaces of a hospital respiratory-isolation facility, Brazil.J Clin Virol. v. 66:62-5. doi: 10.1016/j.jcv.2015.03.008.

FISCHER, M. G. et al. Giant virus with a remarkable complement of genes infects marine zooplankton. Proceedings of the National Academy of Sciences, v. 107, n. 45, p. 19508–19513, 2010.

FRIDMANN-SIRKIS, Y. et al. Efficiency in Complexity: Composition and Dynamic Nature of Mimivirus Replication Factories. Journal of Virology, v. 90, n. 21, p. 10039–10047, 2016.

IYER, L. M.; ARAVIND, L.; KOONIN, E. V. Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. v. 75, n. 23, p. 11720–11734, 2001.

JEUDY, S. et al. Translation in Giant Viruses: A Unique Mixture of Bacterial and Eukaryotic Termination Schemes. PLoS Genetics, v. 8, n. 12, 2012.

KEREPESI, C.; GROLMUSZ, V. Giant viruses of the Kutch Desert. Archives of Virology, v. 161, n. 3, p. 721–724, 2016.

KEREPESI, C.; GROLMUSZ, V. The Giant Virus Finder discovers an abundance of giant viruses in the Antarctic dry valleys. Archives of Virology, p. 1–6, 2017.

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, v. 7, n. JAN, p. 1–9, 2016.

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.

KEREPESI C, GROLMUSZ V. Giant viruses of the Kutch Desert. Arch Virol. 161:721–4, 2016.

KUZNETSOV, Y. G. et al. Morphogenesis of Mimivirus and Its Viral Factories: an Atomic Force Microscopy Study of Infected Cells. Journal of Virology, v. 87, n. 20, p. 11200–11213, 2013.

LA SCOLA, B. et al. A giant virus in amoebae. Science, v. 299, n. 5615, p. 2033, 2003.

LA SCOLA B, DESNUES C, PAGNIER I, ROBERT C, BARRASSI L, FOURNOUS G, MERCHAT M, SUZAN-MONTI M, FORTERRE P, KOONIN E, RAOULT D. The virophage as a unique parasite of the giant mimivirus. Nature 455: 100–104, 2008.

LA SCOLA, B. et al. Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology, v. 53, n. 5, p. 344–353, 2010.

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, v. 112, n. 38, p. E5327–E5335, 2015.

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, v. 111, n. 11, p. 4274–4279, 2014.

LEGENDRE, M. et al. Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. Virology Journal, v.8, n.99, p1-6, 2011.

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. Diversity and evolution of the emerging Pandoraviridae family. Nature communications, v.9, n. 2285, p. 1-12, 2018.

LEGENDRE, M., ALEMPIC, J., PHILIPPE, N., LARTIGUE, A. Pandoravirus Celtis Illustrates the Microevolution Processes at Work in the Giant Pandoraviridae Genomes. Frontiers in Microbiology, v. 10, p.1-11, 2019.

MARURI-AVIDAL L, DOMI A, WEISBERG AS, MOSS B. Participation of vaccinia virus L2 protein in the formation of crescent membranes and immature virions. J Virol 85:2504 –2511. 2011.

MERCER J, HELENIUS A. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. Science 320:531–535, 2008.

MIHARA, T., KOYANO, H., HINGAMP, P., GRIMSLEY, N., GOTO,N., OGATA, H. Taxon Richness of “Megaviridae” Exceeds those of Bacteria and Archaea in the Ocean. Microbes Environ. v. 33, n.2, p.162-171, 2018.

MUTSAFI, Y. et al. Vaccinia-like cytoplasmic replication of the giant Mimivirus. Proceedings of the National Academy of Sciences of the United States of America, v. 107, n. 13, p. 5978–5982, 2010.

PHILIPPE, N. et al. Pandoraviruses: Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes. Science, v. 341, n. 6143, p. 281–286, 2013.

POPGEORGIEV, N.; BOYER, M. Giant Blood Marseillevirus recovered from asymptomatic blood donors. Journal of Infectious, p. 1–27, 2013.

RAOULT, D. The 1.2-Megabase Genome Sequence of Mimivirus. Science, v. 306, n. 5700, p. 1344–1350, 2004.

REED, L. J.; MUENCH, H. A simple method of estimating fifty per cent endpoints. The American Jounal of Hygiene, v. 27, n. 3, p. 493–497, 1938.

RETENO, D. G. et al. Faustovirus, an Asfarvirus-Related New Lineage of Giant Viruses Infecting Amoebae. Journal of Virology, v. 89, n. 13, p. 6585–6594, 2015.

RODRIGUES, R. A. L. et al. Mimivirus Fibrils Are Important for Viral Attachment to the Microbial World by a Diverse Glycoside Interaction Repertoire. Journal of Virology, v. 89, n. 23, p. 11812–11819, 2015.

ROYLES J, AMESBURY MJ, CONVEY P, GRIFFITHS H, HODGSON DA, LENG MJ, CHARMAN DJ. Plants and soil microbes respond to recent warming on the Antarctic Peninsula. Current Biology, 23, 1702–1706, 2013.

SAADI, H.; PAGNIER, I.; et al. First isolation of Mimivirus in a patient with pneumonia. Clinical Infectious Diseases, v. 57, n. 4, p. 127–134, 2013.

SCHEID, P. et al. An extraordinary endocytobiont in Acanthamoeba sp. isolated from a patient with keratitis. Parasitology Research, v. 102, n. 5, p. 945–950, 2008.

SCHULZ, F. et al. Giant viruses with an expanded complement of translation system components. Science, v. 356, n. 6333, p. 82–85, 2017.

SCHULZ, F., ALTEIO, L., GOUDEAU, D., RYAN, E.M., YU, F. B., MALMSTROM, R. R., BLANCHARD, J., WOYKE, T. Hidden diversity of soil giant viruses. Nature Communications, v. 9, n: 4881, p.1-9, 2018.

SHARMA, V.; COLSON, P.; CHABROL, O.; SCHEID, P.; et al. Welcome to pandoraviruses at the “Fourth TRUC” club. Frontiers in Microbiology, v. 6, n. MAY, 2016.

SHEID P. Free-Living Amoebae as Human Parasites and Hosts for Pathogenic Microorganisms. Proceedings, v. 2, n.692, 1-11, 2018.

SUZAN-MONTI, M. et al. Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga Mimivirus. PLoS ONE, v. 2, n. 3, 2007.

SUAREZ C, WELSCH S, CHLANDA P, HAGEN W, HOPPE S, KOLOVOU A, PAGNIER I, RAOULT D, KRIJNSE LOCKER J. Open membranes are the precursors for assembly of large DNA viruses. Cell Microbiol 15:1883–1895, 2013.

SOTO-ARREDONDO, K. J. et al. Biochemical and cellular mechanisms regulating Acanthamoeba castellanii adherence to host cells. Parasitology, v. 141, n. 4, p. 531–541, 2014.

TANVEER T., HAMEED A., GUL A., MATIN A. Quick survey for detection, identification and characterization of Acanthamoeba genotypes from some selected soil and water samples in Pakistan. Annals of Agricultural and Environmental Medicine,v.22, n.2, p.227–230, 2015.

TEMMAM S, DAVOUST B, CHABER AL, LIGNEREUX Y, MICHELLE C, MONTEIL-BOUCHARD S, RAOULT D, DESNUES C. Screening for viral pathogens in African simian bushmeat seized at a French airport. Transbound Emerg Dis 64:1159 – 1167, 2017.

XIAO, C. et al. Structural studies of the giant Mimivirus. PLoS Biology, v. 7, n. 4, p. 0958–0966, 2009.

YOSHIKAWA, G., BLANC-MATHIEU, R., SONG, C., KAYAMA, Y., MOCHIZUKI, T., MURATA, K., OGATA, H., TAKEMUR, M. Medusavirus, a novel large DNA virus discovered from hot spring water. Journal of virology, 2019. Doi: 10.1128/JVI.02130-18.

YOOSUF, N. et al. Related giant viruses in distant locations and different habitats: Acanthamoeba polyphaga moumouvirus represents a third lineage of the Mimiviridae that is close to the Megavirus lineage. Genome Biology and Evolution, v. 4, n. 12, p. 1324–1330, 2012.

YAU S, LAURO FM, MZ DM, BROWN MV, THOMAS T, RAFTERY MJ, et al. Virophage control of antarctic algal host-virus dynamics. Proc Natl Acad Sci.;108:6163–8, 2011.

YUTIN, N.; WOLF, Y. I.; KOONIN, E. V. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. Virology, v. 466–467, p. 38–52, 2014.

ZHANG, W. et al. Four novel algal virus genomes discovered from Yellowstone lake metagenomes. Scientific Reports, v. 5, p. 15131, 2015.

8. OUTRAS ATIVIDADES DESENVOLVIDAS DURANTE O DOUTORADO

8.1 Participação de eventos científicos

- XXVI Congresso Brasileiro de Virologia e V Encontro de Virologia do Mercosul. Período:11 a 14 de outubro de 2015. Florianópolis, SC, Brasil.

- II Simpósio de Microbiologia da UFMG - Microbiologia translacional: do ambiente natural às aplicações biotecnológicas. Período: 05 e 06 de outubro de 2015. Belo Horizonte, MG, Brasil.

- III Simpósio de Microbiologia da UFMG - Doenças microbianas emergentes. Período: 05 e 06 de setembro de 2016. Belo Horizonte, MG, Brasil.

- IV Simpósio de Microbiologia da UFMG - Metabolismo Microbiano: Saúde, Ambiente e Biotecnologia. Período: 02 e 03 de outubro de 2017. Belo Horizonte, MG, Brasil.

- V Simpósio de Microbiologia da UFMG – Desafios atuais no enfrentamento de doenças microbianas. Período: 05 e 06 de setembro de 2018. Belo Horizonte, MG, Brasil.

- XXVIII Congresso Brasileiro de Virologia, VII Encontro de Virologia do Mercosul. Período:06 a 10 de setembro de 2017- **Apresentação oral**

8.2 Supervisão de alunos de iniciação científica

Aluna: Talita Bastos Machado, estudante de graduação de Ciências Biológicas na UFMG. Supervisão no seu projeto de iniciação científica intitulado: Isolamento de vírus gigantes em biomas brasileiros.

8.3 Artigos completos publicados em periódicos indexados durante o período do doutorado (Anexos).

1. Oliveira, G. P.; Andrade, A. C. dos S. P.; Rodrigues, R. L. A.; Arantes, T. S.; Boratto, P. V. M.; Silva, L. S.; Dornas, F. P.; Trindade, G. S.; Drumond, B.; La Scola, B.; Kroon, E. G.; Abrahão, J. S. Promoter Motifs in NCLDVs: An Evolutionary Perspective. Viruses, v.9, p.16 - 20, 2017 DOI: 10.3390/v9010016.

2. Aherfi, S.; Andreani, J.; Baptiste, E.; Oumessoum, A.; Dornas, F. P.; Andrade, A. C. dos S. P.; Chabriere, E.; Abrahao, J.; Levasseur, A.; Raoult, D.; La Scola, B.; Colson, P. A Large Open Pangenome and a Small Core Genome for Giant Pandoraviruses. Frontiers in Microbiology, v. 9, p. 1-13, 2018. DOI: 10.3389/fmicb.2018.01486

3. Rodrigues, R. A. L.; Andreani, J.; Andrade, A. C. dos 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. Journal of Virology, v. 92, p. 1-13, 2018. DOI: 10.1128/JVI.00372-18

4. Silva, L. K. dos S.; Andrade, A. C. dos S. P.; Dornas, F. P.; Rodrigues, R. A. L.; Arantes, T. S. ; Kroon, E. G.; Bonjardim, C. A.; Abrahão, J. S. Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. Scientific Reports, v. 8, p. 1-11, 2018. DOI: 10.1038/s41598-018-22398-3

5. Calixto, R. S.; Oliveira, G. P.; Lima, M. T.; Andrade, A. C. dos S. P.; Trindade, G. de S.; Oliveira, D. B.; Kroon, E. G. A model to study autochthonous Group 1 and 2 Brazilian Vaccinia virus co-infections: development of a qPCR tool and study of pathogenesis. Viruses, v.10, p. 1-10, 2018. DOI: 10.3390/v10010015

6. Rodrigues, R. A. L.; Andrade, Ana C. dos S. P.; Boratto, P. V. de M.; Trindade, G. de S.; Kroon, E. G.; Abrahão, J. S. An Anthropocentric View of the Virosphere-Host Relationship. Frontiers in Microbiology. v.8, p.1-11, 2017. DOI: 10.3389/fmicb.2017.01673

7. Lima, M. L.; Andrade, A. C. dos S. P., Oliveira, G. P., Nicoli, J. R.; Martins, F dos S.; Kroon, E. G.; Abrahão, J. S. Virus and microbiota relationships in humans and other mammals: an evolutionary view. Human microbiome journal, 2019. DOI: https://doi.org/10.1016/j.humic.2018.11.001

8. Boratto, P.; Andrade, A. C. dos P.; Rodrigues, R. A. L.; La Scola, B.; Abrahão, J. The multiple origins of proteins present in tupanvirus particles. Current Opinion in Virology, v. 36, p.25-31, 2019. DOI: https://doi.org/10.1016/j.coviro.2019.02.007

8.4 Trabalhos aceitos para publicação

Capítulo de livro:

Rodrigues, R. A.L., Andrade, A.C.S.P., OLIVEIRA, G. P. Giant viruses and their parasites.Encyclopedia of Virology, 4th edition, 2019.

8.5 Disciplinas cursadas e aproveitamento

Período	Disciplina	Conceito	Créditos
2016/1	Dispensa de crédito*	A	26
2016/2	Tópicos especiais em microbiologia	A	03
2016/2	Microrganismos patogênicos	A	03
2017/1	Tópicos especiais em microbiologia	A	03

*Disciplinas concluídas durante o Mestrado em Microbiologia pela UFMG: Virologia básica, vírus com importância na saúde humana, técnicas básicas em virologia, bacteriologia de anaeróbios, bacteriologia de aeróbios, biologia molecular de micro-organismos, biossegurança e bioética, biologia de leveduras, imunidade inata a infecções microbianas, antimicrobianos, treinamento didático em microbiologia.

ANEXOS

viruses

MDPI

Review

Promoter Motifs in NCLDVs: An Evolutionary Perspective

Graziele Pereira Oliveira [1,†], Ana Cláudia dos Santos Pereira Andrade [1,†], Rodrigo Araújo Lima Rodrigues [1], Thalita Souza Arantes [1], Paulo Victor Miranda Boratto [1], Ludmila Karen dos Santos Silva [1], Fábio Pio Dornas [1], Giliane de Souza Trindade [1], Betânia Paiva Drumond [1], Bernard La Scola [2], Erna Geessien Kroon [1] and Jônatas Santos Abrahão [1,*]

1 Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Minas Gerais, Brazil; graziufmg@yahoo.com.br (G.P.O.); ana.andrade2008@hotmail.com (A.C.d.S.P.A.); rodriguesral07@gmail.com (R.A.L.R.); tsarantes@gmail.com (T.S.A.); pvboratto@gmail.com (P.V.M.B.); ludmilakaren@gmail.com (L.K.d.S.S.); fabiopiod154@gmail.com (F.P.D.); gitrindade@yahoo.com.br (G.d.S.T.); betaniadrumond@gmail.com (B.P.D.); ernagkroon@gmail.com (E.G.K.)

2 Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE) UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université., 27 Boulevard Jean Moulin, Faculté de Médecine, 13385 Marseille Cedex 05, France; bernard.la-scola@univ-amu.fr

* Correspondence: jonatas.abrahao@gmail.com; Tel.: +55-031-3409-3002

† These authors contributed equally to this work.

Academic Editor: Eric Freed

Received: 14 November 2016; Accepted: 5 January 2017; Published: 20 January 2017

Abstract: For many years, gene expression in the three cellular domains has been studied in an attempt to discover sequences associated with the regulation of the transcription process. Some specific transcriptional features were described in viruses, although few studies have been devoted to understanding the evolutionary aspects related to the spread of promoter motifs through related viral families. The discovery of giant viruses and the proposition of the new viral order Megavirales that comprise a monophyletic group, named nucleo-cytoplasmic large DNA viruses (NCLDV), raised new questions in the field. Some putative promoter sequences have already been described for some NCLDV members, bringing new insights into the evolutionary history of these complex microorganisms. In this review, we summarize the main aspects of the transcription regulation process in the three domains of life, followed by a systematic description of what is currently known about promoter regions in several NCLDVs. We also discuss how the analysis of the promoter sequences could bring new ideas about the giant viruses' evolution. Finally, considering a possible common ancestor for the NCLDV group, we discussed possible promoters' evolutionary scenarios and propose the term "MEGA-box" to designate an ancestor promoter motif ('TATATAAAATTGA') that could be evolved gradually by nucleotides' gain and loss and point mutations.

Keywords: megavirales; NCLDV; giant viruses; promoter; transcription; evolution; MEGA-box

1. Introduction

For decades, viruses have been strictly considered intracellular parasites, filterable in membranes of 0.22 nm, composed by genomes of DNA or RNA encoding only a few proteins, being entirely dependent on the metabolic machinery of the host cell [1]. However, viruses show a large diversity of genome size and organization, capsid architecture, mechanisms of replication, and interactions with host cells. The extreme diversity of viruses suggests that they must have had multiple evolutionary origins, thus being polyphyletic [2]. In 2001, a supposedly monophyletic

group named nucleo-cytoplasmic large DNA viruses (NCLDV) was proposed, composed of families *Poxviridae*, *Asfarviridae*, *Iridoviridae* and *Phycodnaviridae* [3]. This group gained notoriety two years later with the discovery of *Acanthamoeba polyphaga mimivirus* [4] and it is currently composed of the families mentioned above, as well as *Ascoviridae*, and the more recently incorporated *Mimiviridae* and *Marseilleviridae* [5]. Moreover, other recently discovered giant viruses such as pandoraviruses, faustoviruses and pithoviruses were classified as members of the NCLDV group [6–9]. This group has single features such as large genomes and a diverse gene repertoire, which encode some proteins never identified previously in viruses. Therefore, the creation of a new viral order named 'Megavirales', encompassing all families of the NCLDV group was proposed [5].

This proposed order comprises viruses with large double-stranded DNA (dsDNA) genomes, encoding hundreds of proteins and capable of infecting a wide-range of eukaryotic organisms. These viruses replicate completely or partly, in the cytoplasm of eukaryotic cells and some of them are able to synthesize RNA polymerases (RNA pol), helicases and transcription factors involved in the transcription initiation and elongation steps with lower dependence of the host's transcriptional machinery [3]. The presence of a robust transcriptional apparatus in some Megavirales members, along with a quasi-autonomous glycosylation and translational machinery, especially in mimiviruses, boosted the discussion about the origin and evolution of giant viruses and their genome. Recent evolutionary reconstructions mapped about 25–50 genes encoding essential genes for the probable most recent common ancestor [10]. Concerning the origin of such giant genomes, different hypotheses have been proposed. Some authors suggest a "genome degradation hypothesis", wherein the giant viruses are derived from a cellular ancestor through genome simplification linked to the adaptation to some host lineage [11,12]. Other authors argue in favor of a "genome expansion hypothesis", wherein the giant viruses evolved from a smaller viral ancestor and the universal genes have been independently acquired from their eukaryotic hosts by progressive gene accretion and duplication. According to this theory, the genes of giant viruses have several origins and the origin of giant viruses is probably from a simpler ancestor [13,14].

On the other hand, the accordion-like model of evolution proposes that there is no trend of genome expansion or general tendency of genome contraction. Instead, viruses evolving by constant gene gain and loss originated from an ancestor giant virus [10]. All these theories are often contradictory and have stimulated discussion about the establishment of a fourth domain of life where the giant viruses of the proposed order Megavirales were suggested to share a common ancestral origin based on analyses of their sequences and gene repertoires and compose a new domain aside Bacteria, Archaea and Eukarya [14–16].

During the last years, a huge effort has been made to better understand the virus–host interaction on many levels. One of the most interesting research fields is how the viruses can explore host transcriptional machinery to express their genes. Nevertheless, it is important also to look into the transcription process of the cellular organisms. The upstream regions of eukaryotes and prokaryotes genes have been studied in different organisms in an attempt to discover sequences associated with the regulation of the transcription process. The same has been done for viruses, especially considering the proposed Megavirales order, where some putative promoter sequences have already been described. In this review, we summarize the main aspects of the transcription regulation process in the three domains of life, followed by a systematic description of current knowledge of the promoter regions of all members within Megavirales order. Finally, we discuss how the analysis of the promoter sequences found in giant viruses provides new insights into the evolutionary history of these complex and intriguing agents.

2. Gene Expression in Cells

In all cells, thousands of genes encoded in the DNA are transcribed into RNA and for the efficient occurrence of this process, multiple events must be triggered. In eukaryotes, the genome is coupled to histones and other proteins, forming the chromatin compact complex. Since wrapping

DNA around histones blocks the access to the genetic information, decondensation of DNA is required, to allow physical access to the the gene locus and the transcription initiation machinery formation [17–19]. The transcription initiation machinery is formed over a region of the genome, the promoter. The promoter is typically located 40 bp upstream and downstream of the transcription start of a gene, called transcription start sites (TSS). Several transcription factors mediate the transcription machinery assembly on the promoter region. There are thousands of transcription factors involved in the transcription process, such as TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH that recognize and bind the promoter region, called the core promoter, and recruit RNA polymerase (RNA pol) [20]. Eukaryotes have five types of RNA pol (I to V). RNA pol I transcribes ribosomal RNA, whereas the type II is the best characterized one and responsible for transcribing genes encoding proteins, and several noncoding RNA classes [18,21,22]. RNA pol III transcribes genes encoding short, untranslated RNAs, such as tRNAs, 5S ribosomal RNA (rRNA) and the spliceosomal U6 small nuclear RNA (snRNA) [23]. RNA pol IV and V transcribe siRNA in plants [24].

One classical element of the core promoter is the TATA-box, which is a consensus sequence (TATAAAT) located at −25 to −30 bp upstream of the TSS. Although the TATA-box sequence is a well-known promoter core motif, it is present only in a minority of mammalian promoters. This sequence is commonly associated with tissue-specific gene transcription and high conservation within species [25,26]. Other eukariotic promoter elements are Initiator (Inr); Downstream Promoter Element (DPE), Core Element Downstream (CED), TFIIB-Recognition Element (TRE), and Motif Ten Element (MTE) [20,27,28]. Together, these components act synergistically to increase transcription efficiency by providing recognition sites for transcription factors, and indicate the direction of transcription and also the DNA strand to be transcribed [20]. The transcription starts with the binding of the TFIID to the TATA-box region, the Inr sequence and/or other core promoter elements [27]. TFIID is a multiprotein complex comprising the TATA-box binding protein (TBP) and more than 10 different TBP associated factors (TAFs) [22]. After binding TBP to the TATA-box motif, the RNA pol II is recruited, and the transcription is triggered (Figure 1A).

Nevertheless, the transcription in eukaryotes is a much more complex process than previously thought and various strategies are used to increase the diversity of transcripts produced. Among mammals, previous analysis has shown that a large proportion of protein-coding genes (58%) use alternative promoters during transcription [25]. These alternative promoters may have different combinations of core promoter elements to increase the variability of transcripts [20,29,30].

There are many differences between the transcription process of eukaryotic and bacteria cells. The bacterial transcription is much simpler compared to the eukaryotic process since the transcription occurs using a single type of RNA pol and there are no transcription factors [31]. This enzyme is capable of synthesizing RNA from a DNA template, but it is unable to locate the promoter and transcription initiation site. Thus, a key factor to transcription is the free subunit named σ (sigma), which is responsible for recognizing the promoter region (Figure 1B) [32,33]. Although the majority of nucleotides within bacteria promoters vary in sequence, several short motifs are conserved. These include the hexamer (TATAAT), located 10 base pairs (bp) upstream of the TSS and is recognized by domain 2 of RNA pol σ subunit. Another motif is the the hexamer (TTGACA), located 35 base pairs (bp) upstream of the TSS and recognized by domain 4 of the RNA pol σ subunit [31,34,35]. In Archaea, there is a mix of eukarya and bacteria translational apparatus. Just as in eukaryotes, the archaea RNA pol is not able to recognize promoter sequences by itself and at least two transcription factors analogous to TBP and TFIIB are required [36–38]. The archeal TBP also recognizes specifically an AT-rich sequence, homologous to the TATA-box region of eukaryotes [39,40]. Although archaea transcription machinery is similar to that of eukaryotes, the characterization of transcription regulators of some archaeas showed that most of the transcriptional regulation in archaea is done by "bacterial-like" regulators, as two homologues of bacterial leucine-responsive regulatory protein (Lrp)—Lrs14 and Sa-Lrp and metal-dependent repressor 1 (MDR1) homologous to bacterial metal-dependent regulators (Figure 1C) [41–43].

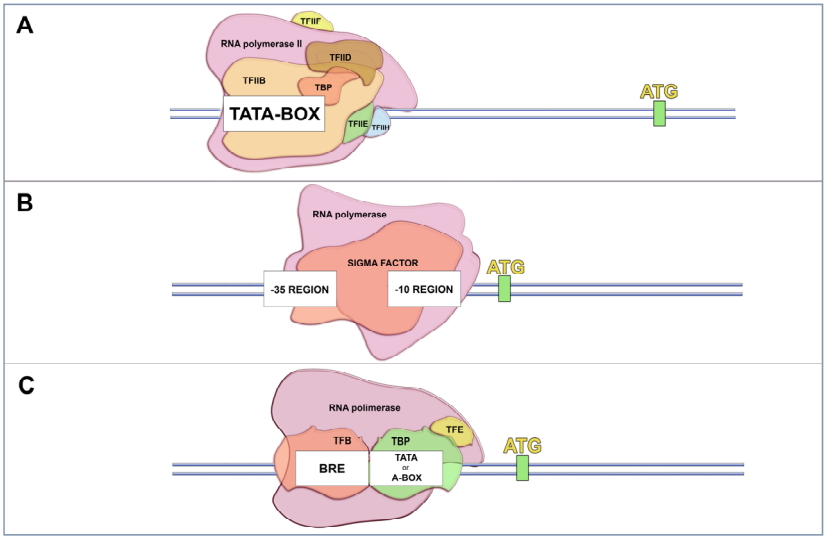


Figure 1. Main features in the transcription initiation machinery presented in the canonical Domains of Life. (**A**) In Eukarya, several components, called general transcription factors (represented as TFIIB, TFIID, TFIIE, TFIIF, TFIIH and TBP), are responsible for assembling over a region called the promoter, where they recruit an RNA polimerase to initiate the transcription process. A classical promoter presented in this group is the TATA-BOX region, located at the positions −25 and −30 from the initial transcription site; (**B**) In Bacteria, the sigma factor recognizes and recruits the RNA polimerase over the promoter regions. These regions are well conserved over the positions −35 and −10 upstream of the initial transcription site; (**C**) Archaea present a mixture of the transcription apparatus of the two other Domains. While the machinery itself is similar to that found in eukaryotes (the general transcription factors, a homologous TATA-BOX region and the RNA polimerase), the archaeal transcription regulators, activators and repressors are homologous to the bacterial ones.

Hypotheses regarding the evolutionary history of translational machinery among the living organisms have been raised during the last years, but the theme is still under debate [44]. Even considering the most recent proposals, the translational process of viruses remains out of the discussion, basically because these organisms are traditionally excluded from the canonical tree of life. However, this scenario has been changing since the discovery of giant viruses [16]. Therefore, it becomes interesting to examine if NCLDV members share similar transcription initiation strategies that could bring insights about how this correlates to giant viruses' evolution.

Gene Expression in NCLDVs

In contrast to cellular genomes, which are formed by dsDNA, viral genomes show a large diversity genome composition, structures, replication and transcription strategies with great implications in virus biology, as virus–host interactions [45]. The majority of the RNA viruses employ virus-coded specific enzymes (RNA-dependent RNA polymerases) to synthesize and modify their mRNA. DNA viruses showing small and intermediate size genomes such as the parvoviruses, papillomaviruses, and adenoviruses, depend on host-cell enzymes for transcription, including the RNA pol [45]. However, viruses with a large genome such as the giant viruses, mostly encode their transcriptional apparatus, which make them relatively independent from their host transcription machinery [15,46].

The transcription of a typical large DNA virus occurs in a temporal pattern in the host cytoplasm (Figure 2). At the start of infection, a subset of immediate early viral proteins is required for DNA replication and host cell manipulation [47,48]. The early mRNAs also encode enzymes and factors needed for transcription of the intermediate genes. Concomitantly with the expression of intermediate genes, the expression of the early genes is often repressed. Finally, late genes are transcribed, directing the synthesis of structural proteins, non-structural proteins and enzymes present in the mature particle

required for viral assembly [45,48]. The efficient transcription of late mRNA usually depends on intermediate gene products, as well as cellular transcription factors that may differ from those used by the early promoters. The products of the late genes include the immediate early transcription factors, which are packaged along with RNA pol and other enzymes within the virus progeny [47–50].

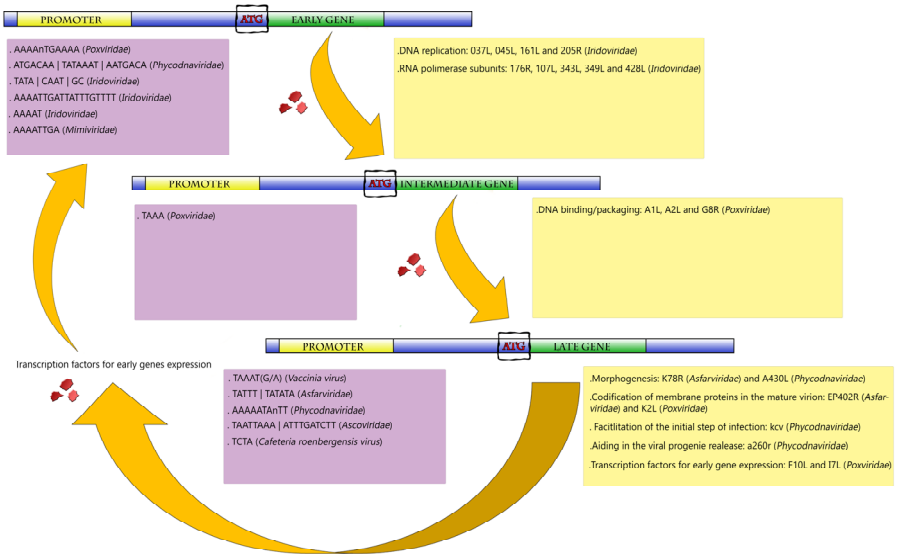


Figure 2. Representative scheme of the temporal gene expression in NCLDVs. During initial times of infection, the expression of genes related to the metabolism of nucleic acids is primarily activated (early and intermediate genes). After DNA replication, the activation of late genes is initiated. Those genes are involved in the production of viral structural proteins, in transcription factors used for early gene expression and also in proteins that facilitate the initial step of infection of the viral progeny in the next round of multiplication. Purple boxes represent the promoters described for giant viruses according to each gene category (early, intermediate and late genes). Yellow boxes exemplify the biological functions involved in each category, with some genes represented inside the parentheses.

This ability to regulate temporally the transcription of genes is characterized as an evolutionary advantage. This strategy is possible due to the presence of promoter codes that dictate when, where, and at what level the classes of early, intermediate, and late genes are transcribed [45,48]. These promoter sequences are different between the three genes classes, but there is a pattern of conservation within the same group. This indicates that during the evolution the gene promoters were selected to ensure the temporal gene expression, and therefore ensure the gene expression in the host cell during its replication [45,47,48,50].

In the following sections, we look closer at how the gene transcription is carried out in each family of the proposed Megavirales order, focusing on the current knowledge about the promoter sequence of these viruses.

3. *Poxviridae* Family

Among NCLDVs, the *Poxviridae* family is one of the most studied. These viruses have enveloped ovoid particles of around 200 nm in diameter and 300 nm in length and present a linear dsDNA genome of approximately 200 kbp coding nearly 200 open reading frames (ORFs). Poxviruses can infect a wide range of hosts, such as insects, birds, and mammals [48,51]. Extensive study of the poxvirus genome and replication cycle allowed a detailed identification of its promoters, as well as important transcription factors. Poxviruses possess their own DNA-dependent RNA polymerase (RNA pol) that is very similar to the eukaryotic protein, regarding size and subunit complexity. In the case of *Vaccinia virus* (VACV), a poxvirus prototype, the enzyme subunits are encoded by eight viral VACV genes

which, in most cases, are homologous to cellular RNApol [52,53]. Gene transcription in poxviruses follows a typical temporal profile regulated by well-conserved promoters of early, intermediate and late genes (Figure 2) [47,48].

The transcription of early genes is characterized by an A/T-rich motif upstream of transcriptional start site with a critical core region located from −13 to −25 to that region. Figure 3 illustrates the promoter motifs described in megavirales members. The representative consensus sequence of the early promoter region is 'AAAANTGAAAA'. Mutagenesis in this promoter region of VACV causes a drastic negative effect on VACV gene transcription [54]. The intermediate genes are transcribed after DNA replication, before the transcription of the late genes. The intermediate core promoter is similar to the early promoter due to the A/T-rich content, but its specific sequence is given by the tetranucleotide 'TAAA'. Furthermore, the intermediate promoter sequence has a bipartite structure presenting a core and an initiator region with similar sequences (TAAA) [55–57]. Three (*A1L*, *A2L*, and *G8R*) of the 53 genes that compose the set of intermediate genes encode transcription factors that are directly related to the late stage of the replication cycle, important to DNA binding/packaging processes and to core-associated proteins [58].

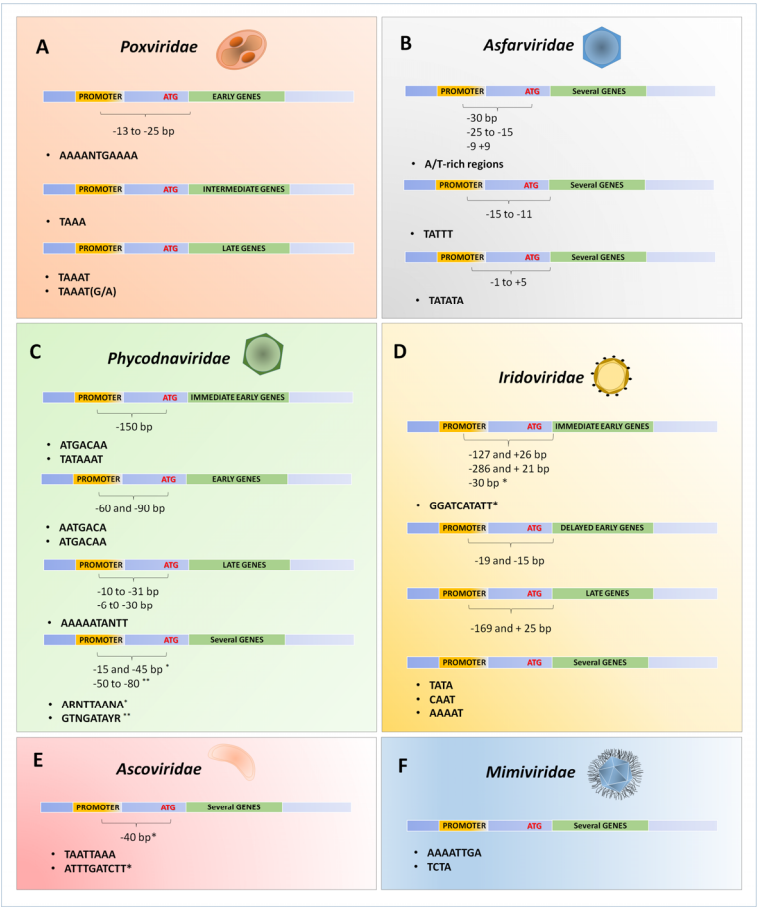


Figure 3. Schematic representation of the promoter's sequences described for different NCLDVs. Compilation of the described promoters for some viral families belonging to the proposed order Megavirales: *Poxviridae* (**A**); *Asfarviridae* (**B**); *Phycodnaviridae* (**C**); *Iridoviridae* (**D**); *Ascoviridae* (**E**) and *Mimiviridae* (**F**). Each promoter was related to the expression of immediate early, early, delayed early, intermediate and late genes, or related to the expression of genes independent of temporal expression (several genes). The distances between the transcription start site or translate start site (ATG) until the promoters are also indicated by brackets.

The transcription of late genes persists until the end of the replication cycle. Around 38 late genes have already been identified, with their main functions related to the codification of membrane proteins in the virion, morphogenesis steps, and also to the production of immediate early transcription factors [57,59]. Most of them are clustered in the central region of the poxviruses genome and also have A/T-rich sequence promoters. These regions consist of a core sequence of about 20 bp with some 'T' residues, separated by a region of about 5–7 bp of a conserved 'TAAAT' motif, which regulates the transcription initiation. Usually, G or A follows the late promoter sequence, performing a 'TAAAT (G/A)' transcription initiation sequence. This sequence is conserved among VACV late promoters, overlapping the site of transcription initiation that is absent in 5' untranslated regions (5'-UTR) [48,54]. Mutations within this conserved element were demonstrated to cause complete inactivation of the promoter, and almost 25% of the 'AAA' sequences are used as transcription initiation sites in VACV. Along with other factors, the viral RNA pol directs the synthesis of late mRNAs, finishing the transcription process [54,60–62].

The presence of complete transcriptional machinery in poxviruses allows a lower dependency of these viruses on their hosts. It permits that the mRNA transcription totally occurs in the host's cytoplasm, right after the virus entry. Addionally, the presence of well conserved promoter regulatory sequences in different poxviruses suggests a conserved evolutionary pattern among them. It is likely that such a complete transcriptional set was already present in their ancestor and was maintained over time. Alternatively, the presence of a robust transcriptional apparatus in all members of the *Poxivirdae* family might be a result of evolutive convergence. Although less parcimonious, the different poxviruses might have had different evolutionary histories regarding the transcription process, including both protein-related elements and promoter sequence regions, but in the course of evolution, they became more similar to each other. It is not yet possible to determine which hypothesis is the correct, or even if other possibilities correspond to the real history of these complex viruses, and this discussion shall continue for a while.

4. *Asfarviridae*

African swine fever virus (ASFV), a large (~200 nm), icosahedral, and enveloped virus is currently the single member of the *Asfarviridae* family, infecting members of the *Suidae* family (pigs, hogs and boars) [63]. The genome is composed of a linear dsDNA molecule of approximately 170 kbp with terminal inverted repeats. It encodes approximately 150 ORFs separated by short intergenic regions [64,65]. ASFV encodes its own RNA pol and all ASFV genes are transcribed by its enzyme [66,67].

Similar to poxviruses, the ASFV gene transcription follows a temporal profile, where immediate early and early genes are expressed before the DNA replication that is followed by the expression of intermediate, late and immediate early genes. Transcription initiation and termination occurs at very precise positions in the genome, encoding a several genes involved in the transcription and modification of viral mRNAs. The transcriptional machinery of ASFV provides an accurate temporal control of gene expression regulated by cis-DNA elements, enhancers, and promoters together with a structural complexity of transcription factors [68]. Analysis of the base composition of the intergenic regions shows that they are rich in A/T sequences, similar to that observed in poxviruses [69–71]. A/T-rich regions located at approximately −30 bp upstream of the ATG translation start site are essential for the expression of the K9L gene, which encodes a protein with similarity to mammalian transcription elongation factor IIS [72]. Furthermore, upstream sequences presented in two intermediate genes exhibit highly conserved sequences at positions −25 to −15, and −9 to +9 to the translational start codon [70]. Experiments involving genetic deletions, linker scan substitutions and point mutations in the promoter sequence of the *p72* gene (major capsid protein) revealed that the replacement of the A/T-rich region by G/C residues strongly reduced the transcription rate, demonstrating the importance of this sequence for efficient late viral transcription [71].

Two other major essential regions for promoter activity are described: one region is located at position −15 to −11 upstream of the transcription start site (TATTT); and the second region at positions −1 to +5 (TATATA) [71]. Mutants presenting the 'TATATA' motif replaced by a G/C-rich sequence had the promoter activity completely abolished, suggesting that ASFV transcription is dependent on such sequence at (or near) the region of transcriptional initiation, similar to what is found in other large viruses [71]. The replacement of the equivalent 'TATATA' sequence on the late genes *K78R*, *EP402R* and *A137R* by the 'GCGC' motif was also demonstrated to be deleterious, suggesting that the A/T-rich sequence could be a motif for late promoter function as well [68,71]. Interestingly, the bipartite structure seen in the late promoter of ASFV is similar to the late and intermediate promoters in poxviruses that contain a core and an initiator region [54,55,62,71]. The similarities found in the transcriptional strategies reinforce the genetic data, indicating a close relationship between poxviruses and asfavirus, pointing to a common ancestor for both viral families.

5. *Phycodnaviridae*

The phycodnaviruses are large and icosahedral viruses (~100–220 nm), with dsDNA genomes ranging from 180 to 560 kbp [73]. Since they infect a diverse group of eukaryotic algae, they are one of the most important groups of organisms regulating the oxygen cycle in the Earth [74,75]. The family *Phycodnaviridae* consists of six genera, named according to the hosts that they infect: *Chlorovirus*, *Coccolithovirus*, *Prasinovirus*, *Prymnesiovirus*, *Phaeovirus*, and *Raphidovirus* [76]. As demonstrated by other giant viruses, the phycodnaviruses exhibit a temporal transcription profile. Early genes are transcribed within 5 to 60 min post-infection (p.i), and transcripts of late genes begin to appear around 60–90 min p.i. However, some early genes can also be detected in later stages of infection [77,78].

The presence of A/T-rich promoters was also observed in phycodnaviruses. Analysis of the *kcv* gene, encoding a potassium ion channel protein in chlorella viruses, revealed a highly conserved 10-nt sequence (AAAAATANTT) in the promoter region of this gene, present in 16 out of 17 chlorellaviruses [77]. This sequence is located at 10–31 nucleotides upstream of the ATG translation start codon in all of the analyzed viruses, and it was associated with late gene transcription, since, apparently, *kcv* transcripts are produced during the late steps of infection. Furthermore, the region that precedes seven genes expressed at later times during the *Paramecium bursaria chlorella virus 1* (PBCV-1) replication cycle (*a85r*, *a237r*, *a248r*, *a260r*, *a292l*, *a430l*, and *a530r*) contain the same sequence or at least a subset of this sequence located at 6–30 nucleotides upstream of the ATG start codon [77]. The study of immediate early genes expressed in chlorovirus infections also revealed A/T-rich sequences as putative promoter regions. Two sequences determined by 'ATGACAA' and 'TATAAAT' (such as the eukaryotic "TATA-box") were located in a 150 bp region from the translation start codon in the upstream regions of almost all immediate early genes (20 of 23 studied) [78]. These elements, especially 'ATGACAA', were absent in all genes so far examined, expressed after 40 min p.i, including *A122R* (Vp260) [79], *A181-182R* (chitinase), *A292L* (chitosanase) [80], *A430L* (major capsid protein) [81], *vAL-1* [82].

Bioinformatics analysis revealed highly conserved nucleotide sequences in putative promoter regions involving three different chlorella viruses: PBCV-1, virus MT325 [83], and *Paramecium bursaria chlorella virus* NY-2A [84]. Three putative AT-rich sequence promoters, comprising seven to nine nucleotides (ARNTTAANA, AATGACA and GTNGATAYR), located at 150-nt upstream of the translation start codon of many ORFs were observed [85]. The 'ARNTTAANA' sequence is found between nucleotides −15 and −45 relative to the ATG translation start codon. This sequence occurs in the promoter region of 25% of PBCV-1 genes, 22% of NY-2A genes and 12% of MT325 genes. Regarding the entire genome, this sequence is present within the 200-nt promoter region during 44% of the time in PBCV-1, 49% of the time in NY-2A, and 37% of the time in MT325. The hotspot for the presence of the 'AATGACA' sequence is located between nucleotides −60 and −90 from the translational start codon. This sequence occurs in the promoter region of 16% of the PBCV-1 genes, 18% of NY-2A genes and 8% of MT325 genes. Regarding the entire genome, this sequence is present within the 200-nt promoter region in 54% of the PBCV-1 genes, 53% of the NY-2A genes, and 25% of the MT325 genes [85].

The 'AATGACA' sequence in PBCV-1 is associated with early genes during the replication cycle [85]. This sequence is very similar to a motif previously identified in some chlorella viruses (ATGACAA), which is also correlated with early transcripts [78]. Finally, the 'GTNGATAYR' sequence is mainly located at nucleotide positions −50 to −80 from the ATG initiation codon, occurring in the promoter region of 13% of PBCV-1 genes, 14% NY-2A genes, and in 11% of MT325 genes. Regarding the entire genome, this sequence is found specifically within the 200-nt promoter region in 28% of the PBCV-1 genes, 22% of the NY-2A genes, and 21% of the MT325 genes [85].

Unlike other members of the NCLDVs, phycodnaviruses do not encode their own RNA pol and need to appropriate the host's RNA pol to properly make their transcripts [86]. However, uniquely for the *Phycodnaviridae* family, *Emiliania huxleyi virus 86* (EhV-86), a coccolithovirus that infects the marine calcifying microalga *Emiliania huxleyi*, contains a total of six RNA pol subunits, which suggests that this virus partially encodes its own transcription machinery [87]. Although these viruses present some important elements for the mRNA synthesis, it is not possible to state that they have their own transcriptional complete apparatus, at least for the majority of them. Therefore, concerning the transcriptional process, the phycodnaviruses seem to present a different evolutionary history.

6. *Iridoviridae*

The *Iridoviridae* family is composed by five genera: *Ranavirus*, *Megalocytivirus* and *Lymphocystivirus* that infect vertebrates; *Iridovirus* and *Chloriridovirus* that infect invertebrates [88]. Iridoviruses have a linear dsDNA genome varying from 105 to 212 kbp, coding between 92 and 211 putative proteins. They present a non-enveloped icosahedral particle of 300 nm in size [89–92]. These large viruses also display a pattern of temporal gene expression regulation, wherein the genes are divided into three classes: immediate-early (IE or α), delayed-early (DE or β), and late (L or γ) genes [93–95]. Iridoviruses are typical nucleo-cytoplasmic viruses. They begin the replication cycle in the nucleus, followed by the second phase of genome replication in the cytoplasm [90].

Gene transcription and promoter sequences studies have been performed for only a few genes in members of the *Iridoviridae* family. The study of promoter sequences in iridovirus is focused mainly in the *Ranavirus* genus (using type species *Frog virus 3* (FV3)) and *Iridovirus* genus (using type species *Invertebrate iridescent virus 6* (IIV-6)), the type species of the *Iridovirus* genus. Notwithstanding, both the gene expression and promoter sequences studies have been performed for only a few genes in the *Iridoviridae* family. The most complex studies were performed with immediate-early *ICR-169* and *ICR-489* genes of FV3 [96,97]. Those studies revealed the importance of a 78 bp sequence before the transcription start site of an IE gene of the FV3 promoter. It was shown that an FV3 protein acts in trans to induce the transcription of the major FV3 IE gene, *ICR-169*, and is dependent on the 78 bp sequence located at the 5′ position from the start site of the transcription of this gene [98]. Two years later, the same group demonstrated that a 23 bp sequence was possibly a critical *cis*-regulatory element for the occurrence of FV3 *trans*-activation, since a significant reduction of transcription occurred after its deletion, located at the 5′ region, showing the sequence 'ATATCTCACAGGGGAATTGAAAC' [96]. Despite the importance of the approximately 23-nt sequence upstream of the transcription start site in the IE *ICR-169* gene of FV3, this sequence had no similarity with the promoter region of the intermediate gene *ICR489*. This lack of similarity indicated that the contemporary regulation of these two promoters is not controlled by sequences upstream of the start point of transcription [97]. It is worthy to note that in the *ICR489* gene, in an upstream region, 'TATA', 'CAAT', and 'GC' motifs were identified, which are similar to those of typical eukaryotic promoters [97].

Another study analyzed three genes—two early (*ICP-18* and *ICP-46*) and a late one [major capsid protein (MCP)] of *Bohle iridovirus* (another *Ranavirus* member)—looking for conserved regions to be considered as regulatory elements [99]. The authors demonstrated that all gene promoters included sequences located 127 to 281 bases upstream of the transcription initiation site (127 pb or ICB-18, 281 pb for ICP46, and 169 pb for MCP), but also sequences located from 21 to 26 bases downstream of this site (26 bases for *ICP-18*, 21 bases for ICP 46 and 25 bases for MCP) [99].

Moreover, a detailed study conducted in the following years identified an essential 'AAAAT' motif in a DE gene of IIV-6 (*Iridovirus*) [100]. The authors described a sequence of 19 bp (AAAATTGATTATTTGTTTT), located between −19 and −2 relative to the mRNA transcription start site, which is the putative region responsible for promoter activity of the DNApol gene. Deletions and point mutations in the DNApol promoter of IIV-6 showed that each of the 5-nt of 'AAAAT' motif located between −19 and −15 were equally essential for promoter activity. Mutations at the downstream side had a lower effect, but the role of individual nucleotides positioned at −14 to −5 was not analyzed in this study [100].

It is noteworthy that the same critical 'AAAAT' motif was found in the 100-nt upstream of the putative translational start codons of several other putative DE IIV-6 genes [91]. In *Invertebrate iridescent virus 3* (IIV-3), many homologues of these genes also presented the 'AAAAT' motif in proximity to their start codon. A great similarity was also found between the region upstream of the DNApol ORF and the corresponding region in 12 iridovirus genomes [101]. Eight of these genomes showed a similar 'AAAAT' motif in the DNApol upstream region and three sequenced ranavirus genomes also shared the related 'TAAAT' motif in their DNA pol promoter region, which may indicate a conserved regulation of DE promoter activity in iridoviruses [101].

A study that targeted a IE gene (012L) of IIV-6 showed that the transcription start site is located 30-nt upstream of the ATG translational start codon. Analyzing mutants (produced by deletion), it was established that the intergenic region located between −21 and −10 (GGATCATATT) upstream of the transcription start site comprised the promoter sequence promoter 012L gene. This type of sequence was not observed in upstream regions of other IE genes of IIV-6, such as 468R, 006L and 010R. The 'TATA' and 'CAAT' sequences were also identified in the intergenic region of this gene, as well as sequences similar to the 'AAAAT' motif described to the DNA pol gene, but this sequence had no promoter activity for the 012L, differently than demonstrated for the DNA pol gene. The 037L and 012L genes of IIV-6, both early genes, do not share conserved key promoter motifs. However, DNA pol is considered a DE gene and 012L an IE gene [100,102].

Despite the presence of homologs of RNA pol subunits in the iridoviruses genome, host RNA pol II is required for the synthesis of *Ranavirus* IE transcripts, and it is likely that the same is true from *Iridovirus* IE genes, contrasting to pox- and asfaviruses [103–106]. It has been proposed that the RNA pol subunits found in members of the *Iridoviridae* family are probably involved in the cytoplasmic phase of transcription in later stages of infection [91,107]. Such a paradox may reflect the long co-evolution period that these viruses had been through. It is possible that the ancestor of iridoviruses presented a complete transcription apparatus, but some elements were lost due to the adaptation to a more parasitic lifestyle. Other possibilities are the occurrence of events of horizontal gene transfer (HGT) between the viruses and their hosts. However, the lack of information about such events involving members of the *Iridoviridae* family prevents further insights into this alternative for the evolution of the transcription apparatus of these viruses.

7. *Ascoviridae*

The *Ascoviridae* family has two genera that include *Ascovirus*, with three species including *Spodoptera frugiperda ascovirus 1a* (SfAV-1a), the prototype of the genus, *Trichoplusia ni ascovirus 2a* (TnAV-2a), and *Heliothis virescens ascovirus 3a* (HvAV-3a), and the *Toursvirus* genus, with only one representative, *Diadromus pulchellus ascovirus 4a* (DpAV-4a) [108,109]. Ascoviruses are enveloped viruses, 300–400 nm long by 100–150 nm in diameter, with a circular dsDNA genome with sizes ranging from 116 to 185 kb, infecting arthropods, mainly lepidopterans [110–112].

The studies regarding the ascoviruses are still in their infancy. Information about the replication and more specifically, the transcription process, are extremely scarce. The current knowledge about transcription in ascoviruses come from the analyses of the *Ascovirus* genus [110,113]. A study performed using a possible variant of HvAV-3, the *Spodoptera exigua ascovirus 5a* (SeAV-5a) showed that the 5′-UTR region of the SeAV-5a MCP gene is composed of 25-nt [114]. The upstream region

of this gene does not present a typical eukaryotic class II promoter motif sequence 'TATAAAT' (TATA box). However, the putative 5' transcription control region of the SeAV-5a MCP gene shares similarities with other ascoviruses and iridoviruses, containing a conserved TATA-box like motif (TAATTAAA) and an 'ATTTGATCTT' motif within 40-nt upstream of the translation initiation codon ATG [114]. The 'TAATTAAA' and 'ATTTGATCTT' motifs are located downstream and upstream of the transcription initiation site, respectively. Furthermore, the ORF p27 presents a similar 5' downstream transcription promoter region, suggesting that such a region might be a truly regulatory sequence within ascoviruses [114].

Sequences from the promoter regions of the MCP genes from ascoviruses and IIV-6 (late genes), showed that ascoviruses and iridoviruses are closely related in this aspect, suggesting that the transcription regulation could be maintained during the viral evolution process in closely related viruses [115,116]. Furthermore, phylogenetic studies showed that ascoviruses probably evolved from the iridoviruses [116–118]. It is possible that the same pattern of temporary gene expression exhibited in iridoviruses (and the other members of proposed Megavirales order) was conserved in the ascoviruses lineage, and that such a mechanism might have been present in their common ancestor.

8. *Mimiviridae* and Other Amoebal Giant Viruses

The discovery of mimiviruses in 2003 and the establishment of the *Mimiviridae* family astonished the scientific community, making the term 'giant virus' more appropriated than ever. These viruses have particles visible in light microscopy, with sizes of ~700 nm in diameter. Viral particles have characteristics never described before in the virosphere, such as long proteic fibrils (~125 nm in length) immersed in a peptidoglycan matrix, and a star-shaped face, named stargate, responsible for the releasing of the genome inside the cytoplasm of their host (*Acanthamoeba* genus) [4,119–121]. The genome is a linear dsDNA molecule of about 1.2 Mbp, coding more than 1000 proteins, including a large set of transcriptional elements [15,122].

Similar to other NCLDVs members, mimiviruses genes can be divided into early, intermediate and late categories according to three major temporal classes of transcription determined by mRNA deep sequencing [49]. The analysis of the intergenic regions of *Acanthamoeba polyphaga mimivirus*, the prototype species of *Mimivirus* genus, showed a conserved 'AAAATTGA' motif in nearly 50% of genes [50]. The intergenic regions of the genome of mimiviruses have an average size of 157-nt. In silico analyses showed that the conserved 'AAAATTGA' motifs are present within the 150-nt upstream regions of the translation start codon in 45% of all predicted mimivirus genes [50]. This motif is mainly associated to early (or the late-early) genes during the viral infectious cycle, and it is absent from the upstream regions of mimivirus late genes, such as DNA replication and particle morphogenesis and assembly. It is noteworthy that similar sequences were described regulating the early genes in other giant viruses, such as iridoviruses and phycodnaviruses, as described in the topics above. Besides the early promoter sequence, another A/T-rich motif (two 10-nt informative segments separated by a highly degenerated 4-nt sequence) was identified as a putative late promoter within mimiviruses, which is present in 24.2% of the considered late class genes. To the best of our knowledge, an intermediate promoter sequence has not already been described in mimiviruses [49,50].

In a distant relative, the *Cafeteria roenbergensis virus* [CroV (*Cafeteria* genus)], *Mimiviridae* family; the same early promoter motif was identified in the upstream region of 35% of genes [123]. However, considering the late promoter motif, this virus exhibits a different putative regulatory sequence compared to other mimiviruses, wherein the 'TCTA' tetramer flanked by A/T-rich regions on either side was found in the 5' upstream of 124 late genes [123]. Moreover, CroV present eight RNA pol II subunits, six transcription factors, several helicases, among others, indicating the presence of nearly complete transcriptional machinery. This feature seems to be a mark to all members of the *Mimiviridae* family, which suggests that such a robust transcriptional apparatus was already present in the last common ancestor.

After the discovery of mimiviruses, other giant viruses infecting amoebae were described, such as marseilleviruses, which is currently classified in the family *Marseilleviridae* [124]. Other viruses have also been isolated but still not properly classified, namely faustoviruses [125], pandoraviruses [8,126], phitoviruses [127,128] and mollivirus [129]. Although these viruses are not yet officially recognized by the ICTV, they are genuine members of the NCLDVs [6,7,9]. In all of these giant viruses, a set of transcriptional elements has already been identified, including many RNA pol subunits, indicating a nearly autonomous process in these viruses. However, analysis of promoters and studies aiming to understand how gene expression is regulated in those newly discovered viruses remain to be performed.

9. MEGA-Box: A Putative Promoter Region in the Common Ancestor of Megavirales

The proposed Megavirales order comprises viral families that exhibit some unique features that allow their clustering into a monophyletic group [5]. In addition to some core genes that are shared among these viruses, they present other similarities, such as a temporal transcription profile. As described above, all viruses present elements to the transcriptional apparatus, most of them reaching up to the independence from their host in this step of the viral life cycle. Also, the presence of an A/T-rich promoter sequence has been described in many representatives of each family, even in those in which the genome presents a high G/C content. More interesting is the fact that some promoter sequences found in one family are very similar to others found in their relatives (Figure 3). This fact suggests that a possible common ancestor of the Megavirales order likely had an A/T-rich promoter sequence. More interesting is the fact that some promoter sequences found in one family are very similar to others found in their giant relatives. This fact suggests that such a common ancestor of Megavirales likely had an A/T-rich promoter sequence.

The origin of the members of the Megavirales order is still under debate, but the evolutionary history of some of its members is already being told, at least concerning genome evolution. The first members to be analyzed were the poxviruses. It has been demonstrated by phylogenetic analysis based on the presence/absence of genes that genomes from this family have been subject to frequent events of gene duplication, deletion, and HGT from their hosts. Many of these genes can interfere with host immune signaling, such as homologues of cytokines receptors which could confer some advantages in the interaction with the hosts [130–132]. By analyzing the poxviruses' closest relative, ASFV, it seems that it has been through the same pattern of evolution, at least considering the multigene and p22 gene families [133,134].

The "accordion-like" pattern of evolution was also identified in different members of the *Iridoviridae* family. It is particularly interesting the fact that iridoviruses infecting the same host-range exhibited a similar pattern of gene gain and loss, but this was slightly different when the viruses infected different hosts (fish vs. insect-infecting viruses), suggesting that such a pattern was driven by host–virus co-evolution [135]. Finally, the same evolutionary model for members of the families *Phycodnaviridae* and *Mimiviridae* has recently been described. The genomic comparisons of closely related viruses belonging to the *Mimiviridae* and *Phycodnaviridae* families show that genomes accumulating genomic mutations occur on successive cycles of genome expansion and reduction. In addition, there is no general tendency of genome expansion or contraction. Each family exhibits a specific pattern for gene acquisition, which might be a reflex of interaction with distinct hosts [10]. Since these viruses seem to exhibit a similar pattern of genome evolution, it is possible that a similar scenario has also happened with their promoter sequences. In the same way, it is reasonable to consider that NCLDVs' common ancestor evolved by the same "accordion-like" pattern, and thus it presented a promoter region that underwent an analogous mechanism.

Considering a common origin for the NCLDVs, a possible scenario is that the Megavirales' common ancestor presented a 'TATATAAAATTGA' promoter motif, which we named here as the "MEGA-box" (an allusion to the conserved TATA-box promoter found in cellular organisms). Over time, with the Megavirales' order radiation, the MEGA-box has been gradually evolved by nucleotides'

gain and loss, analogously to that reported for the entire genome, which evolved through gene gain and loss. The MEGA-box was slightly modified in the poxviruses lineage, at least concerning the early promoter motif. Considering the intermediate and the late promoter motifs of poxviruses, if they truly came from the MEGA-box, this could have happened through a series of nucleotide loss. However, it is also possible that the emergence of other promoters, rather than the early one, have emerged after the establishment of the poxvirus' lineage, thus not originating from the ancestral promoter sequence. The same might be true for mimiviruses, phycodnaviruses and iridoviruses. Considering asfavirus and ascoviruses, their promoter sequences might have originated from the MEGA-box through successive gain and loss of nucleotides. However, another scenario is also possible, wherein their promoter motifs emerged from the poxviruses and iridoviruses lineages respectively (closest evolutionary groups). This scenario is in agreement with the proposition that the Megavirales' ancestor was already a giant virus with a large genome [10]. In this aspect, the giant ancestor also had a large promoter sequence that evolved through constant nucleotide gain and loss, a pattern analogous to the accordion-like model of genome evolution. However, other scenarios are also possible, although less probable, considering the evolutionary data currently available for these viruses. One is that the ancestor had a very short promoter sequence, like a poxvirus intermediate promoter (TAAA), that underwent massive nucleotide gain over time, leading to very large promoter sequences in the majority of the giant viruses. Another one is just the opposite; wherein the ancestor had a very large promoter region that had been losing nucleotides during evolution. A third pathway, equally unlikely, would be the acquisition of promoter sequences by horizontal/lateral transfer. Similar to different genes, the MEGA-box promoter evolutionary pattern during the radiation of NCLDVs members could be related to the co-evolution with different hosts over time.

Whether the NCLDVs came from a simple entity [14,136], or from an already complex organism [10,16,137], is still under debate. Despite this, increasing evidence that they originated from a common ancestor is emerging, and it suggests that such an ancestor evolved through an "accordion-like" pattern. By analyzing the promoter regions currently known for different giant viruses, we provide another piece of evidence to support this statement. Further, we propose how a conserved A/T-rich promoter sequence was present in the possible common ancestor, which might have evolved by continuous gain and loss of nucleotides, in addition to some point mutations in the MEGA-box original sequence. Other scenarios could also be discussed for the evolution of the promoter sequences of the NCLDVs, including selective sweep or convergence. However, these alternatives run off the diffused hypothesis of a common origin for the putative Megavirales order.

10. What Comes Next?

Most of the giant viruses have a powerful genetic arsenal, encoding several proteins necessary for the transcription system which provides a relative independence of their hosts for this process. In addition, the transcription of this high gene content is temporally regulated by promoter regions that exhibit some similarities, indicating a common origin of these regulatory elements. Although many studies have already been done in relation to almost all viral families of the Megavirales order, most of them remain without biological confirmation; i.e., the promoter motifs in many giant viruses were predicted, but not experimentally validated. Therefore, the performance of biological studies to confirm the existence and the effect of all promoter motifs described so far in giant viruses is imperative. This analysis will truly establish the common temporal regulation pattern predicted in these viruses, and will also corroborate (or even refute) the hypothesis of an A/T-rich promoter in the Megavirales common ancestor. Moreover, the deep analysis of the genome of the recently described giant viruses (Marseilleviruses, Pandoraviruses, Pithoviruses, Faustoviruses and Mollivirus), and also the discovery of new complex viruses, will strongly contribute to complete the puzzle of the origin and evolution of Megavirales.

On the other hand, the biotechnology field will also be boosted by the advance in the studies of promoters and gene expression in giant viruses. Among the NCLDVs, the poxviruses are by far the best

characterized group regarding the genome expression, especially the VACV. These viruses have been used as expression vectors for the synthesis of proteins and as vaccine candidates to prevent infectious diseases and treat cancer, mainly due to their high gene expression levels [69,138]. This attribute is clearly shared with other giant viruses that were recently described, and the real comprehension of their gene regulation and expression will bring uncountable possibilities for biotechnology purposes. Finally, the impact of the giant viruses on the basic comprehension of the origin and evolution of life is undeniable, as well as for their ecological, medical and technological importance. The discovery of even more complex viruses associated with the advance of many techniques used for genomic studies will certainly answer those remaining questions around the NCLDVs, and will surely bring new exciting challenges for the whole scientific community.

Acknowledgments: We would like to thank our colleagues from GEPVIG (Grupo de Estudo e Prospecção de Vírus Gigantes) and Laboratório de Vírus of Universidade Federal de Minas Gerais. In addition, 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). G.S.T., E.G.K. and J.S.A. are CNPq researchers. B.L.S., J.S.A. and E.G.K. are members of a CAPES-COFECUB project.

Author Contributions: G.P.O.; A.C.S.P.A.; R.A.L.R.; T.S.A.; P.V.M.B.; L.K.S.S. and F.P.D. wrote the paper. B.P.D., G.S.T., B.L.S., E.G.K. and J.S.A. approved and reviewed the final version.

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

References

1. Lwoff, A. The concept of virus. *J. Gen. Microbiol.* **1957**, *17*, 239–253. [CrossRef] [PubMed]
2. Durzyńska, J.; Goździcka-Józefiak, A. Viruses and cells intertwined since the dawn of evolution. *Virol. J.* **2015**, *16*, 169. [CrossRef] [PubMed]
3. Iyer, L.M.; Aravind, L.; Koonin, E.V. Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.* **2001**, *75*, 11720. [CrossRef] [PubMed]
4. 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*. [CrossRef] [PubMed]
5. Colson, P.; de Lamballerie, X.; Fournous, G.; Raoult, D. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. *Intervirology* **2012**, *55*, 321–332. [CrossRef] [PubMed]
6. Sharma, V.; Colson, P.; Chabrol, O.; Pontarotti, P.; Raoult, D. *Pithovirus sibericum*, a new bona fide member of the "Fourth TRUC" club. *Front. Microbiol.* **2015**, *4*, 722. [CrossRef] [PubMed]
7. Sharma, V.; Colson, P.; Chabrol, O.; Scheid, P.; Pontarotti, P.; Raoult, D. Welcome to pandoraviruses at the "Fourth TRUC" club. *Front. Microbiol.* **2015**, *18*, 423.
8. Scheid, P. A strange endocytobiont revealed as largest virus. *Curr. Opin. Microbiol.* **2016**, *31*, 58–62. [CrossRef] [PubMed]
9. Benamar, S.; Reteno, D.G.; Bandaly, V.; Labas, N.; Raoult, D.; la Scola, B. Faustoviruses: Comparative genomics of new Megavirales family members. *Front. Microbiol.* **2016**, *5*, 3. [CrossRef] [PubMed]
10. Filée, J. Genomic comparison of closely related Giant Viruses supports an accordion-like model of evolution. *Front. Microbiol.* **2015**, *16*, 593. [CrossRef] [PubMed]
11. Suzan-Monti, M.; la Scola, B.; Raoult, D. Genomic and evolutionary aspects of Mimivirus. *Virus Res.* **2006**, *117*, 145–155. [CrossRef] [PubMed]
12. Claverie, J.M. Viruses take center stage in cellular evolution. *Genome Biol.* **2006**, *7*, 110. [CrossRef] [PubMed]
13. Moreira, D.; López-García, P. Comment on "The 1.2-megabase genome sequence of Mimivirus". *Science* **2005**, *20*, 1114. [CrossRef] [PubMed]
14. Yutin, N.; Wolf, Y.I.; Koonin, E.V. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* **2014**, *38–52*, 466–467. [CrossRef] [PubMed]
15. Raoult, D.; Audic, S.; Robert, C.; Abergel, C.; Renesto, P.; Ogata, H.; la Scola, B.; Suzan, M.; Claverie, J.M. The 1.2-megabase genome sequence of Mimivirus. *Science* **2004**, *306*, 1344–1350. [CrossRef] [PubMed]
16. Boyer, M.; Madoui, M.A.; Gimenez, G.; La Scola, B.; Raoult, D. Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. *PLoS ONE* **2010**, *2*, e15530. [CrossRef] [PubMed]

17. Schones, D.E.; Cui, K.; Cuddapah, S.; Roh, T.-Y.; Barski, A.; Wang, Z.; Wei, G.; Zhao, K. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **2008**, *132*, 887–898. [CrossRef] [PubMed]
18. Fuda, N.J.; Ardehali, M.B.; Lis, J.T. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **2009**, *10*, 186–192. [CrossRef] [PubMed]
19. Lacadie, S.A.; Ibrahim, M.M.; Gokhale, S.A.; Ohler, U. Divergent transcription and epigenetic directionality of human promoters. *FEBS J.* **2016**, *283*, 4214–4222. [CrossRef] [PubMed]
20. Haberle, V.; Lenhard, B. Promoter architectures and developmental gene regulation. *Semin. Cell Dev. Biol.* **2016**, *57*, 11–23. [CrossRef] [PubMed]
21. Weiss, S.; Gladstone, L. A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. *J. Am. Chem. Soc.* **1959**, *81*, 4118–4119. [CrossRef]
22. Thomas, M.C.; Chiang, C.M. The general transcription machinery and generalcofactors. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41*, 105–178. [CrossRef] [PubMed]
23. Sabin, L.R.; Delás, M.J.; Hannon, G.J. Dogma derailed: The many influences of RNA on the genome. *Mol. Cell* **2013**, *49*, 783–794. [CrossRef] [PubMed]
24. Vannini, A.; Ringel, R.; Kusser, A.G.; Berninghausen, O.; Kassavetis, G.A.; Cramer, P. Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell* **2010**, *143*, 59–70. [CrossRef] [PubMed]
25. Carninci, P.; Sandelin, A.; Lenhard, B.; Katayama, S.; Shimokawa, K.; Ponjavic, J.; Semple, C.A.; Taylor, M.S.; Engström, P.G.; Frith, M.C. Genome-wide analysis of mammalian promoter architecture andevolution. *Nat. Genet.* **2006**, *38*, 626–635. [CrossRef] [PubMed]
26. Cooper, S.J.; Trinklein, N.D.; Anton, E.D.; Nguyen, L.; Myers, R.M. Comprehensive analysis of transcriptional promoter structure and functionin 1% of the human genome. *Genome Res.* **2006**, *16*, 1–10. [CrossRef] [PubMed]
27. Smale, S.T.; Kadonaga, J.T. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **2003**, *72*, 449–479. [CrossRef] [PubMed]
28. Maston, G.A.; Evans, S.K.; Green, M.R. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genom. Hum. Genet.* **2006**, *7*, 29–59. [CrossRef] [PubMed]
29. Hansen, S.K.; Tjian, R. TAFs and TFIIA mediate differential utilization of thetandem Adh promoters. *Cell* **1995**, *82*, 565–575. [CrossRef]
30. Ren, B.; Maniatis, T. Regulation of Drosophila Adh promoter switching by aninitiator-targeted repression mechanism. *EMBO J.* **1998**, *17*, 1076–1086. [CrossRef] [PubMed]
31. Browning, D.F.; Busby, S.J. Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **2016**, *14*, 638–650. [CrossRef] [PubMed]
32. Burgess, R.R.; Travers, A.A.; Dunn, J.J.; Bautz, E.K. Factor stimulating transcription by RNA polymerase. *Nature* **1969**, *221*, 43–46. [CrossRef] [PubMed]
33. Lee, D.J.; Minchin, S.D.; Busby, S.J. Activating transcription in bacteria. *Annu. Rev. Microbiol.* **2012**, *66*, 125–152. [CrossRef] [PubMed]
34. Murakami, K.S.; Masuda, S.; Campbell, E.A.; Muzzin, O.; Darst, S.A. Structural basis of transcription initiation: An RNA polymerase holoenzyme–DNA complex. *Science* **2002**, *296*, 1285–1290. [CrossRef] [PubMed]
35. Browning, D.F.; Busby, S.J. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2004**, *2*, 57–65. [CrossRef] [PubMed]
36. Hausner, W.; Wettach, J.; Hethke, C.; Thomm, M. Two transcription factors related with the eucaryal transcription factors TATA binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J. Biol. Chem.* **1996**, *271*, 30144–30148. [CrossRef] [PubMed]
37. Bell, S.D.; Jaxel, C.; Nadal, M.; Kosa, P.F.; Jackson, S.P. Temperature template topology, and factor requirements of archaeal transcription. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15218–15222. [CrossRef] [PubMed]
38. Darcy, T.J.; Hausner, W.; Awery, D.E.; Edwards, A.M.; Thomm, M.; Reeve, J.N. Methanobacterium thermoautotrophicum RNA polymerase and transcription in vitro. *J. Bacteriol.* **1999**, *181*, 4424–4429. [PubMed]
39. Reiter, W.D.; Hudepohl, U.; Zillig, W. Mutational analysis of an archae bacterial promoter—Essential role of a TATA Box for transcription efficiency and start-site selection in vitro. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9509–9513. [CrossRef] [PubMed]

40. Palmer, J.R.; Daniels, C.J. In vivo definition of an archaeal promoter. *J. Bact.* **1995**, *177*, 1844–1849. [CrossRef] [PubMed]
41. Kyrpides, N.C.; Ouzounis, C.A. Transcription in Archaea. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8545–8550. [CrossRef] [PubMed]
42. Aravind, L.; Koonin, E.V. DNA-binding proteins and evolution of transcription regulation in the *Archaea*. *Nucleic Acids Res.* **1999**, *27*, 4658–4670. [CrossRef] [PubMed]
43. Bell, S.D.; Jackson, S.P. Mechanism and regulation of transcription in *Archaea*. *Curr. Opin. Microbiol.* **2001**, *4*, 208–213. [CrossRef]
44. Werner, F.; Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **2011**, *9*, 85–98. [CrossRef] [PubMed]
45. Whelan, S. Viral Replication Strategies. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, USA, 2014; Volume 2, p. 2160.
46. Abergel, C.; Rudinger-Thirion, J.; Giege, R.; Claverie, J.M. Virus-encoded aminoacyl-tRNA synthetases: Structural and functional characterization of mimivirus TyrRS and MetRS. *J. Virol.* **2007**, *81*, 12406–12417. [CrossRef] [PubMed]
47. Broyles, S.S.; Knutson, B.A. Poxvirus transcription. *Future Virol.* **2010**, *5*, 639–650. [CrossRef]
48. Moss, B. Poxviridae. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, USA, 2014; Volume 2, p. 2129.
49. Legendre, M.; Audic, S.; Poirot, O.; Hingamp, P.; Seltzer, V.; Byrne, D.; Lartigue, A.; Lescot, M.; Bernadac, A.; Poulain, J.; et al. mRNA deep sequencing reveals 75 new genes and a complex transcriptional landscape in Mimivirus. *Genome Res.* **2010**, *20*, 664–674. [CrossRef] [PubMed]
50. Suhre, K.; Audic, S.; Claverie, J.M. Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14689–14693. [CrossRef] [PubMed]
51. Damon, I.K. Poxviruses. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, USA, 2014; Volume 2, p. 2160.
52. Baroudy, B.M.; Moss, B. Purification and characterization of a DNA dependent RNA polymerase from vaccinia virions. *J. Biol. Chem.* **1980**, *255*, 4372–4380. [PubMed]
53. Knutson, B.A.; Broyles, S.S. Expansion of poxvirus RNA polymerase subunits sharing homology with corresponding subunits of RNA polymerase II. *Virus Genes* **2008**, *36*, 307–311. [CrossRef] [PubMed]
54. Davison, A.J.; Moss, B. The structure of vaccinia virus early promoters. *J. Mol. Biol.* **1989**, *210*, 749–769. [CrossRef]
55. Baldick, C.J.; Keck, J.G.; Moss, B. Mutational analysis of the core, spacer and initiator regions of vaccinia virus intermediate class promoters. *J. Virol.* **1992**, *66*, 4710–4719. [PubMed]
56. Knutson, B.A.; Liu, X.; Oh, J. Vaccinia virus intermediate and late promoter elements are targeted by the TATA-binding protein. *J. Virol.* **2006**, *80*, 6784–6793. [CrossRef] [PubMed]
57. Yang, Z.; Bruno, D.P.; Martens, C.A.; Porcella, S.F.; Moss, B. Genome-wide analysis of the 5′ and 3′ ends of vaccinia virus early mRNAs delineates regulatory sequences of annotated and anomalous transcripts. *J. Virol.* **2011**, *85*, 5897–5909. [CrossRef] [PubMed]
58. Keck, J.G.; Baldick, C.J.; Moss, B. Role of DNA replication in vaccinia virus gene expression: A naked template is required for transcription of three late transactivator genes. *Cell* **1990**, *61*, 801–809. [CrossRef]
59. Yang, Z.; Reynolds, S.E.; Martens, C.A.; Bruno, D.P.; Porcella, S.F.; Moss, B. Expression profiling of the intermediate and late stages of poxvirus replication. *J. Virol.* **2011**, *85*, 9899–9908. [CrossRef] [PubMed]
60. Yang, Z.; Martens, C.A.; Bruno, D.P.; Porcella, S.F.; Moss, B. Pervasive initiation and 3′-end formation of poxvirus postreplicative RNAs. *J. Biol. Chem.* **2012**, *287*, 31050–31060. [CrossRef] [PubMed]
61. Broyles, S.S.; Liu, X.; Zhu, M.; Kremer, M. Transcription factor YY1 is a vaccinia virus late promoter activator. *J. Biol. Chem.* **1999**, *274*, 35662–35667. [CrossRef] [PubMed]
62. Hänggi, M.; Bannwarth, W.; Stunnenberg, H.G. Conserved TAAAT motif in vaccinia virus late promoters: Overlapping TATA box and site of transcription initiation. *EMBO J.* **1986**, *5*, 1071–1076. [PubMed]
63. International Committee on Taxonomy of Viruses. Available online: http://www.ictvonline.org/taxonomyHistory.asp?taxnode_id=20151927&taxa_name=Asfarviridae (accessed on 22 September 2016).
64. Sogo, J.M.; Almendral, J.M.; Talavera, A.; Vinuela, E. Terminal and internal inverted repetitions in African swine fever virus DNA. *Virology* **1984**, *133*, 271–275. [CrossRef]

65. Tulman, E.R.; Delhon, G.A.; Ku, B.K.; Rock, D.L. African swine fever virus. *Curr. Top. Microbiol. Immunol.* **2009**, *328*, 43–87. [PubMed]
66. Kuznar, J.; Salas, M.L.; Vinuela, E. DNA-dependent RNA polymerase in African swine fever virus. *Virology* **1980**, *101*, 169–175. [CrossRef]
67. Salas, J.; Salas, M.L.; Vinuela, E. Effect of inhibitors of the host cell RNA polymerase II on African swine fever virus multiplication. *Virology* **1988**, *164*, 280–283. [CrossRef]
68. Rodríguez, J.M.; Salas, M.L. African swine fever virus transcription. *Virus Res.* **2013**, *173*, 15–28. [CrossRef] [PubMed]
69. Moss, B. Genetically engineered poxviruses for recombinant gene expression, vaccination, and safety. *Proc. Natl. Acad. Sci. USA* **1996**, *15*, 11341–11348. [CrossRef]
70. Rodríguez, J.M.; Salas, M.L.; Vinuela, E. Intermediate class of mRNAs in African swine fever virus. *J. Virol.* **1996**, *70*, 8584–8589. [PubMed]
71. Garcia-Escudero, R.; Vinuela, E. Structure of African swine fever virus late promoters: Requirement of a TATA sequence at the initiation region. *J. Virol.* **2000**, *74*, 8176–8182. [CrossRef] [PubMed]
72. Yates, P.R.; Dixon, L.K.; Turner, P.C. Promoter analysis of an African swine fever virus gene encoding a putative elongation factor. *Biochem. Soc. Trans.* **1995**, *23*, 139S. [CrossRef]
73. Van Etten, J.L.; Graves, M.V.; Müller, D.G.; Boland, W.; Delaroque, N. Phycodnaviridae—Large DNA algal viruses. *Arch. Virol.* **2002**, *147*, 1479–1516. [CrossRef] [PubMed]
74. Van Etten, J.L.; Meints, R.H. Giant viruses infecting algae. *Annu. Rev. Microbiol.* **1999**, *53*, 447–494. [CrossRef] [PubMed]
75. Wilson, W.H.; van Etten, J.L.; Allen, M.J. The *Phycodnaviridae*: The story of how tiny giants rule the world. *Curr. Top. Microbiol. Immunol.* **2009**, *328*, 1–42. [PubMed]
76. International Committee on Taxonomy of Viruses. Available online: http://www.ictvonline.org/taxonomyHistory.asp?taxnode_id=20153552&taxa_name=Phycodnaviridae (accessed on 22 September 2016).
77. Kang, M.; Graves, M.; Mehmel, M.; Moroni, A.; Gazzarrini, S.; Thiel, G.; Gurnon, J.R.; van Etten, J.L. Genetic diversity in chlorella viruses flanking *kcv*, a gene that encodes a potassium ion channel protein. *Virology* **2004**, *326*, 150–159. [CrossRef] [PubMed]
78. Kawasaki, T.; Tanaka, M.; Fujie, M.; Usami, S.; Yamada, T. Immediate early genes expressed in chlorovirus infections. *Virology* **2004**, *318*, 214–223. [CrossRef] [PubMed]
79. Chuchird, N.; Nishida, K.; Kawasaki, T.; Fujie, M.; Usami, S.; Yamada, T. A variable region on the chlorovirus CVK2 genome contains five copies of the gene for Vp260, a viral-surface glycoprotein. *Virology* **2002**, *10*, 289–298. [CrossRef] [PubMed]
80. Yamada, T.; Hiramatsu, S.; Songsri, P.; Fujie, M. Alternative expression of a chitosanase gene produces two different proteins in cells infected with Chlorella virus CVK2. *Virology* **1997**, *14*, 361–368. [CrossRef] [PubMed]
81. Graves, M.V.; Meints, R.H. Characterization of the major capsid protein and cloning of its gene from algal virus PBCV-1. *Virology* **1992**, *188*, 198–207. [CrossRef]
82. Sugimoto, I.; Hiramatsu, S.; Murakami, D.; Fujie, M.; Usami, S.; Yamada, T. Algal-lytic activities encoded by Chlorella virus CVK2. *Virology* **2000**, *10*, 119–126. [CrossRef] [PubMed]
83. Fitzgerald, L.A.; Graves, M.V.; Li, X.; Feldblyum, T.; Hartigan, J.; van Etten, J.L. Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect Chlorella Pbi. *Virology* **2007**, *20*, 459–471. [CrossRef] [PubMed]
84. Fitzgerald, L.A.; Graves, M.V.; Li, X.; Feldblyum, T.; Nierman, W.C.; van Etten, J.L. Sequence and annotation of the 369-kb NY-2A and the 345-kb AR158 viruses that infect Chlorella NC64A. *Virology* **2007**, *20*, 472–484. [CrossRef] [PubMed]
85. Fitzgerald, L.A.; Boucher, P.T.; Yanai-Balser, G.M.; Suhre, K.; Graves, M.V.; van Etten, J.L. Putative gene promoter sequences in the chlorella viruses. *Virology* **2008**, *25*, 388–393. [CrossRef] [PubMed]
86. Van Etten, J.L. Unusual life style of giant chlorella viruses. *Ann. Rev. Genet.* **2003**, *37*, 153–195. [CrossRef] [PubMed]
87. Wilson, W.H.; Schroeder, D.C.; Allen, M.J.; Holden, M.T.; Parkhill, J.; Barrell, B.G.; Churcher, C.; Hamlin, N.; Mungall, K.; Norbertczak, H.; et al. Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* **2005**, *12*, 1090–1092. [CrossRef] [PubMed]

88. International Committee on Taxonomy of Viruses. Available online: http://www.ictvonline.org/taxonomyHistory.asp.taxnode_id=20153003&taxa_name=Iridoviridae (accessed on 22 September 2016).
89. Darai, G.; Delius, H.; Clarke, J.; Apfel, H.; Schnitzler, P.; Flugel, R.M. Molecular cloning and physical mapping of the genome of fish lymphocystis disease virus. *Virology* **1985**, *146*, 292–301. [CrossRef]
90. Williams, T.; Barbosa-Solomieu, V.; Chinchar, V.G. A decade of advances in Iridovirus research. *Adv. Virus Res.* **2005**, *65*, 174–248.
91. Jakob, N.J.; Muller, K.; Bahr, U.; Darai, G. Analysis of the first complete DNA sequence of an invertebrate iridovirus: Coding strategy of the genome of Chilo iridescent virus. *Virology* **2001**, *286*, 182–196. [CrossRef] [PubMed]
92. Chinchar, V.G.; Yu, K.H.; Jancovich, J.K. The molecular biology of frog virus 3 and other iridoviruses infecting cold-blooded vertebrates. *Viruses* **2011**, *3*, 1959–1985. [CrossRef] [PubMed]
93. Barray, S.; Devauchelle, G. Protein synthesis in cells infected by Chilo iridescent virus: Evidence for temporal control of three classes of induced polypeptides. *Virology* **1987**, *138*, 253–261.
94. D'Costa, S.M.; Yao, H.; Bilimoria, S.L. Transcription and temporal cascade in Chilo iridescent virus infected cells. *Arch. Virol.* **2001**, *146*, 2165–2178. [CrossRef] [PubMed]
95. D'Costa, S.M.; Yao, H.J.; Bilimoria, S.L. Transcriptional mapping in Chilo iridescent virus infections. *Arch. Virol.* **2004**, *149*, 723–742. [CrossRef] [PubMed]
96. Willis, D.B. DNA sequences required for trans activation of an immediate- early frog virus 3 gene. *Virology* **1987**, *161*, 1–7. [CrossRef]
97. Beckman, W.; Tham, T.N.; Aubertin, A.M.; Willis, D.B. Structure and regulation of the immediate-early frog virus 3 gene that encodes ICR489. *J. Virol.* **1988**, *62*, 1271–1277. [PubMed]
98. Willis, D.B.; Granoff, A. Transactivation of an immediateearly frog virus 3 promoter by a virion protein. *J. Virol.* **1985**, *56*, 495–501. [PubMed]
99. Pallister, J.; Goldie, S.; Coupar, B.; Hyatt, A. Promoter activity in the 59 flanking regions of the Bohle iridovirus ICP 18, ICP 46 and major capsid protein genes. *Arch. Virol.* **2005**, *150*, 1911–1919. [CrossRef] [PubMed]
100. Nalcacioglu, R.; Ince, I.A.; Vlak, J.M.; Demirbag, Z.; van Oers, M.M. The Chilo iridescent virus DNA polymerase promoter contains an essential AAAAT motif. *J. Gen. Virol.* **2007**, *88*, 2488–2494. [CrossRef] [PubMed]
101. Nalcacioglu, R.; Demirbag, Z.; Vlak, J.M.; van Oers, M.M. Promoter analysis of the Chilo iridescent virus DNA polymerase and major capsid protein genes. *Virology* **2003**, *317*, 321–329. [CrossRef] [PubMed]
102. Dizman, Y.A.; Demirbag, Z.; Ince, I.A.; Nalcacioglu, R. Transcriptomic analysis of Chilo iridescent virus immediate early promoter. *Virus Res.* **2012**, *167*, 353–357. [CrossRef] [PubMed]
103. Goorha, R. Frog virus 3 requires RNA polymerase II for its replication. *J. Virol.* **1981**, *37*, 496–499. [PubMed]
104. Goorha, R.; Willis, D.B.; Granoff, A. Macromolecular synthesis in cells infected by frog virus 3. VI. Frog virus 3 replication is dependent on the cell nucleus. *J. Virol.* **1977**, *21*, 802–805. [PubMed]
105. Goorha, R.; Murti, G.; Granoff, A.; Tirey, R. Macromolecular synthesis in cells infected by frog virus 3. VIII. The nucleus is a site of frog virus 3 DNA and RNA synthesis. *Virology* **1978**, *84*, 32–50. [CrossRef]
106. Baroudy, B.M.; Moss, B. Sequence homologies of diverse length tandem repetitions near ends of vaccinia virus genome suggest unequal crossing over. *Nucleic Acids Res.* **1982**, *25*, 5673–5679. [CrossRef]
107. Tidona, C.A.; Darai, G. Molecular anatomy of lymphocystis disease virus. *Arch. Virol. Suppl.* **1997**, *13*, 49–56. [PubMed]
108. Federici, B.A.; Bigot, Y.; Granados, R.R.; Hamm, J.J.; Miller, L.K.; Newton, I.; Stasiak, K.; Vlak, J.M. Family Ascoviridae. In *Virus Taxonomy: 8th Report of the International Committee on Taxonomy of Viruses*; Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A., Eds.; Elsevier Academic Press: San Diego, CA, USA, 2005; pp. 261–265.
109. International Committee on Taxonomy of Viruses. Available online: http://www.ictvonline.org/taxonomyHistory.asp?taxnode_id=20151919&taxa_name=Ascoviridae (accessed on 22 September 2016).
110. Cheng, X.W.; Wang, L.; Carner, G.R.; Arif, B.M. Characterization of three ascovirus isolates from cotton insects. *J. Invertebr. Pathol.* **2005**, *89*, 193–202. [CrossRef] [PubMed]
111. Asgari, S.; Davis, J.; Wood, D.; Wilson, P.; McGrath, A. Sequence and organization of the Heliothis virescens ascovirus genome. *J. Gen. Virol.* **2007**, *88*, 1120–1132. [CrossRef] [PubMed]

112. Bigot, Y.; Rabouille, A.; Sizaret, P.Y.; Hamelin, M.H.; Periquet, G. Particle and genomic characterization of a new member of the Ascoviridae, *Diadromus pulchellus ascovirus*. *J. Gen. Virol.* **1997**, *78*, 1139–1147. [CrossRef] [PubMed]
113. Cheng, X.W.; Carner, G.R.; Arif, B.M. A new ascovirus from *Spodoptera exigua* and its relatedness to the isolate from *Spodoptera frugiperda*. *J. Gen. Virol.* **2000**, *81*, 3083–3092. [CrossRef] [PubMed]
114. Salem, T.Z.; Turney, C.M.; Wang, L.; Xue, J.; Wan, X.-F.; Cheng, X.-W. Transcriptional analysis of a major capsid protein gene from *Spodoptera exigua ascovirus 5a*. *Arch. Virol.* **2008**, *153*, 149–162. [CrossRef] [PubMed]
115. Zhao, K.; Cui, L.W. Molecular characterization of the major virion protein gene from the *Trichoplusia ni ascovirus*. *Virus Genes* **2003**, *27*, 93–102. [CrossRef] [PubMed]
116. Stasiak, K.; Renault, S.; Demattei, M.V.; Bigot, Y.; Federici, B.A. Evidence for the evolution of ascoviruses from iridoviruses. *J. Gen. Virol.* **2003**, *84*, 2999–3009. [CrossRef] [PubMed]
117. Chinchar, V.G.; Hyatt, A.; Miyazaki, T.; Williams, T. Family *Iridoviridae*: Poor viral relations no longer. *Curr. Top. Microbiol. Immunol.* **2009**, *328*, 123–170. [PubMed]
118. Piégu, B.; Asgari, S.; Bideshi, D.; Federici, B.A.; Bigot, Y. Evolutionary relationships of iridoviruses and divergence of ascoviruses from invertebrate iridoviruses in the superfamily Megavirales. *Mol. Phylogenet. Evol.* **2015**, *84*, 44–52. [CrossRef] [PubMed]
119. 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**, *13*, 114. [CrossRef] [PubMed]
120. Xiao, C.; Kuznetsov, Y.G.; Sun, S.; Hafenstein, S.L.; Kostyuchenko, V.A.; Chipman, P.R.; Suzan-Monti, M.; Raoult, D.; McPherson, A.; Rossmann, M.G. Structural studies of the giant mimivirus. *PLoS Biol.* **2009**, *28*, e92. [CrossRef] [PubMed]
121. Rodrigues, R.A.; dos Santos Silva, L.K.; Dornas, F.P.; de Oliveira, D.B.; Magalhães, T.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] [PubMed]
122. Legendre, M.; Santini, S.; Rico, A.; Abergel, C.; Claverie, J.M. Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. *Virol. J.* **2011**, *4*, 8–99. [CrossRef] [PubMed]
123. 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* **2010**, *9*, 19508–19513. [CrossRef] [PubMed]
124. Colson, P.; de Lamballerie, X.; Yutin, N.; Asgari, S.; Bigot, Y.; Bideshi, D.K.; Cheng, X.W.; Federici, B.A.; van Etten, J.L.; Koonin, E.V.; et al. "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* **2013**, *158*, 2517–2521. [CrossRef] [PubMed]
125. 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]
126. 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] [PubMed]
127. 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. USA* **2014**, *111*, 4274–4279. [CrossRef] [PubMed]
128. Levasseur, A.; Andreani, J.; Delerce, J.; Bou Khalil, J.; Robert, C.; La Scola, B.; Raoult, D. Comparison of a modern and fossil pithovirus reveals its genetic conservation and evolution. *Genome Biol. Evol.* **2016**, *25*, 2333–2339. [CrossRef] [PubMed]
129. 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*, 5327–5335. [CrossRef] [PubMed]
130. McLysaght, A.; Baldi, P.F.; Gaut, B.S. Extensive gene gain associated with adaptive evolution of poxviruses. *Proc. Natl. Acad. Sci. USA* **2003**, *23*, 15655–15660. [CrossRef] [PubMed]
131. Hughes, A.L.; Friedman, R. Poxvirus genome evolution by gene gain and loss. *Mol. Phylogenet. Evol.* **2005**, *35*, 186–195. [CrossRef] [PubMed]

132. Elde, N.C.; Child, S.J.; Eickbush, M.T.; Kitzman, J.O.; Rogers, K.S.; Shendure, J.; Geballe, A.P.; Malik, H.S. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* **2012**, *17*, 831–841. [CrossRef] [PubMed]
133. Dixon, L.K.; Chapman, D.A.; Netherton, C.L.; Upton, C. African swine fever virus replication and genomics. *Virus Res.* **2013**, *173*, 3–14. [CrossRef] [PubMed]
134. Portugal, R.; Coelho, J.; Höper, D.; Little, N.S.; Smithson, C.; Upton, C.; Martins, C.; Leitão, A.; Keil, G.M. Related strains of African swine fever virus with different virulence: Genome comparison and analysis. *J. Gen. Virol.* **2015**, *96*, 408–419. [CrossRef] [PubMed]
135. Huang, Y.; Huang, X.; Liu, H.; Gong, J.; Ouyang, Z.; Cui, H.; Cao, J.; Zhao, Y.; Wang, X.; Jiang, Y.; et al. Complete sequence determination of a novel reptile iridovirus isolated from soft-shelled turtle and evolutionary analysis of Iridoviridae. *BMC Genomics* **2009**, *14*, 224. [CrossRef] [PubMed]
136. Koonin, E.V.; Krupovic, M.; Yutin, N. Evolution of double-stranded DNA viruses of eukaryotes: From bacteriophages to transposons to giant viruses. *Ann. N. Y. Acad. Sci.* **2015**, *1341*, 10–24. [CrossRef] [PubMed]
137. Nasir, A.; Kim, K.M.; Caetano-Anolles, G. Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evol. Biol.* **2012**, *24*. [CrossRef] [PubMed]
138. Moss, B. Reflections on the early development of poxvirus vectors 2013. *Vaccine* **2013**, *6*, 4220–4222. [CrossRef] [PubMed]

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 July 2018
doi: 10.3389/fmicb.2018.01486

A Large Open Pangenome and a Small Core Genome for Giant Pandoraviruses

Sarah Aherfi[1], Julien Andreani[1], Emeline Baptiste[1], Amina Oumessoum[1], Fábio P. Dornas[2], Ana Claudia dos S. P. Andrade[2], Eric Chabriere[1], Jonatas Abrahao[2], Anthony Levasseur[1], Didier Raoult[1], Bernard La Scola[1] and Philippe Colson[1]**

[1] Microbes Evolution Phylogenie et Infections (MEϕI), Institut Hospitalo-Universitaire Méditerranée Infection, Assistance Publique – Hôpitaux de Marseille, Institut de Recherche pour le Développement, Aix-Marseille Université, Marseille, France, [2] Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

OPEN ACCESS

Edited by:
Sead Sabanadzovic, Mississippi State University, United States

Reviewed by:
Luis Carlos Guimarães, Universidade Federal do Pará, Brazil
Subir Sarker, La Trobe University, Australia

***Correspondence:**
Bernard La Scola
bernard.la-scola@univ-amu.fr
Philippe Colson
philippe.colson@univ-amu.fr

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *21 February 2018*
Accepted: *14 June 2018*
Published: *10 July 2018*

Citation:
Aherfi S, Andreani J, Baptiste E, Oumessoum A, Dornas FP, Andrade ACSP, Chabriere E, Abrahao J, Levasseur A, Raoult D, La Scola B and Colson P (2018) A Large Open Pangenome and a Small Core Genome for Giant Pandoraviruses. Front. Microbiol. 9:1486. doi: 10.3389/fmicb.2018.01486

Giant viruses of amoebae are distinct from classical viruses by the giant size of their virions and genomes. Pandoraviruses are the record holders in size of genomes and number of predicted genes. Three strains, *P. salinus*, *P. dulcis*, and *P. inopinatum*, have been described to date. We isolated three new ones, namely *P. massiliensis*, *P. braziliensis*, and *P. pampulha*, from environmental samples collected in Brazil. We describe here their genomes, the transcriptome and proteome of *P. massiliensis*, and the pangenome of the group encompassing the six pandoravirus isolates. Genome sequencing was performed with an Illumina MiSeq instrument. Genome annotation was performed using GeneMarkS and Prodigal softwares and comparative genomic analyses. The core genome and pangenome were determined using notably ProteinOrtho and CD-HIT programs. Transcriptomics was performed for *P. massiliensis* with the Illumina MiSeq instrument; proteomics was also performed for this virus using 1D/2D gel electrophoresis and mass spectrometry on a Synapt G2Si Q-TOF traveling wave mobility spectrometer. The genomes of the three new pandoraviruses are comprised between 1.6 and 1.8 Mbp. The genomes of *P. massiliensis*, *P. pampulha*, and *P. braziliensis* were predicted to harbor 1,414, 2,368, and 2,696 genes, respectively. These genes comprise up to 67% of ORFans. Phylogenomic analyses showed that *P. massiliensis* and *P. braziliensis* were more closely related to each other than to the other pandoraviruses. The core genome of pandoraviruses comprises 352 clusters of genes, and the ratio core genome/pangenome is less than 0.05. The extinction curve shows clearly that the pangenome is still open. A quarter of the gene content of *P. massiliensis* was detected by transcriptomics. In addition, a product for a total of 162 open reading frames were found by proteomic analysis of *P. massiliensis* virions, including notably the products of 28 ORFans, 99 hypothetical proteins, and 90 core genes. Further analyses should allow to gain a better knowledge and understanding of the evolution and origin of these giant pandoraviruses, and of their relationships with viruses and cellular microorganisms.

Keywords: pandoravirus, giant virus, megavirales, pangenome, core genome

INTRODUCTION

Giant viruses of amoebae are distinct from classical viruses by many features, primarily by the giant size of their virions and genomes (Colson et al., 2017a). The first to be discovered was Mimivirus, in 2003 (La Scola et al., 2003). Since then, giant viruses that were described were classified into two viral families and several new putative viral groups (Colson et al., 2017b). Their remarkable characteristics and expanding diversity have raised many questions about their origin and evolution. Notably, these giant viruses display several traits that are hallmarks of cellular organisms, including the encoding of several translation components by their genomes. Pandoraviruses were discovered in 2013 (Philippe et al., 2013). The first pandoravirus was isolated from a marine sediment layer of a river on a coast of Chile (Philippe et al., 2013), the second one from a freshwater pond in Australia (Philippe et al., 2013), and the third one from contact lenses and their storage case fluid of a keratitis patient in Germany (Scheid et al., 2014). These viruses hence appear to be cosmopolitan, and pandoravirus-like sequences were detected in metagenomes generated from water and soil samples collected worldwide (Verneau et al., 2016; Kerepesi and Grolmusz, 2017; Brinkman et al., 2018) as well as from mosquitoes (Temmam et al., 2015; Atoni et al., 2018), biting midges (Temmam et al., 2015), and simian bushmeat and human plasma (Verneau et al., 2016; Temmam et al., 2017). Pandoraviruses became, and still are, the record holders in size of viral genomes and number of predicted genes. In addition, their virions exhibit a weird morphology for viruses, being ovoid, surrounded by a tegument-resembling structure, and devoid of recognizable capsid (Philippe et al., 2013). As for the mimiviruses, they had been for years mingled with intra-amoebal eukaryotic parasites (Scheid et al., 2014).

The isolation of all giant viruses of amoebae until now was made possible through the use of amoebae of the genus *Acanthamoeba* or *Vermamoeba* as culture support (Khalil et al., 2017). This culture strategy has been considerably optimized during the past 15 years, with, notably, the implementation of high-throughput amoebal co-culture protocols (Khalil et al., 2017). Such approach was recently used to discover new giant viruses of amoebae in Brazil (Dornas et al., 2015). Consequently, three new pandoraviruses were isolated in 2015–2016 from water collected from a Soda lake and from soil samples (Dornas et al., 2015). We describe here the genomes of these three new giant viruses and the pangenome of pandoraviruses based on these three new isolates and the three previously described strains, namely *Pandoravirus dulcis*, *Pandoravirus salinus* (Philippe et al., 2013), and *Pandoravirus inopinatum* (Scheid, 2016).

MATERIALS AND METHODS

Virus Isolation, Production, and Purification

After collection, samples were stored at −80°C and then co-cultured on *Acanthamoeba castellanii*, as previously described (Andreani et al., 2016). The three samples induced amoebal lysis, and then were subcultured to produce the new virus isolates. Viruses were then purified and concentrated by centrifugation (Andreani et al., 2016).

Genome Sequencing

The viral genomes were sequenced on the Illumina MiSeq instrument (Illumina, Inc., San Diego, CA, United States) by using both paired-end and mate-pair strategies for *P. massiliensis* and *P. braziliensis*, and paired-end strategy only for *P. pampulha*. Genomic DNA was quantified by a Qubit assay with the high-sensitivity kit (Life technologies, Carlsbad, CA, United States). DNA paired-end libraries were constructed with 1 ng of each genome as input with the Nextera XT DNA sample prep kit (Illumina, Inc., San Diego, CA, United States), according to the manufacturer's recommendations. Automated cluster generation and paired-end sequencing with dual index reads were performed in a single 39-h run in 2 × 250 bp. Paired-end reads were trimmed and filtered according to read qualities. The mate-pair library was prepared with 1.5 μg of genomic DNA. Genomic DNA was simultaneously fragmented and tagged with a mate-pair junction adapter. The library profile and the concentration were visualized on a high-sensitivity bioanalyzer labchip (Agilent Technologies Inc., Santa Clara, CA, United States). In each construction, libraries were normalized at 2 nM and pooled, denaturated, and diluted to reach a concentration of 15 pM, before being loaded onto the reagent cartridge, then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-h run generating 2 × 151-bp long reads. The quality of the genomic data was analyzed by FastQC[1].

Genome Assembly

The three pandoravirus genomes were assembled using CLC genomics v.7.5[2] with default parameters. The assembly of the *P. massiliensis* genome provided nine scaffolds. Gaps were filled and scaffolds were then reordered using both Sanger sequencing and three different assembly tools used in combination, including A5, Velvet, and ABySS (Simpson et al., 2009; Zerbino, 2010; Tritt et al., 2012). The genome of *P. braziliensis* was assembled into seven scaffolds, which were then reordered into two scaffolds by using similarity searches and synteny bloc detection with the closest available genomes. Long-range PCR was performed to resolve the linear or circular organization of the two scaffolds. The genome of *P. pampulha* was assembled into 45 scaffolds that were reordered and fused to form one fragment, using the same strategy than for *P. braziliensis*.

Transcriptome Sequencing of *Pandoravirus massiliensis*

The transcriptome of *P. massiliensis* was analyzed at the following times: 30 min (t0), then 2 (t2h), 4 (t4h), 6 (t6h), and 8 h (t8h) after inoculation of the virus on *A. castellanii* in Peptone Yeast Glucose browth medium. At each time point, the co-culture was centrifuged then immediately frozen at −80°C. RNA was extracted with the RNeasy mini kit (Qiagen, Hilden, Germany).

[1]https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[2]https://www.qiagenbioinformatics.com/

After cDNA generation by RT-PCR, libraries were constructed with the Nextera XT DNA sample prep kit. cDNA was quantified by a Qubit assay with the high-sensitivity kit. To prepare the paired-end library, dilution was performed to require 1 ng of each genome as input. The "tagmentation" step fragmented and tagged the DNA. Then, limited cycle PCR amplification (12 cycles) completed tag adapters and introduced dual-index barcodes. The library profile was validated on an Agilent 2100 Bioanalyzer with a DNA high-sensitivity labchip (Agilent Technologies Inc., Santa Clara, CA, United States), and the fragment size was estimated to be 1.5 kbp. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, United States), libraries were normalized on specific beads according to the Nextera XT protocol (Illumina, Inc.). Normalized libraries were pooled for sequencing on the MiSeq instrument. Automated cluster generation and paired-end sequencing with dual index reads were performed in a single 39-h run in 2 × 250 bp. Total information of 3.6 Gb was obtained from a 370 k/mm^2 cluster density with a cluster passing quality control filters of 95.7% (6,901,000 passed filtered clusters). Within this run, the index representation for *P. massiliensis* infection kinetic was respectively determined to be 2.5, 9.4, 3.7, 15.1, and 0.6%. Finally, paired-end reads were trimmed and filtered according to the read qualities.

Proteome Analysis of *Pandoravirus massiliensis*

Preparation of the Total Proteins of the Virus

Samples were rapidly lysed in DTT solubilization buffer (2% SDS, 40 mM Tris–HCl, pH 8.0, 60 mM DTT) with brief sonication. The 2D Clean-Up kit eliminated nucleic acids, salts, lipids, and other reagents not compatible with immunoelectrophoresis.

Two-Dimensional Gels

Analysis of the 1D gel electrophoresis was performed with the Ettan IPGphor II control software (GE Healthcare). For the 2D gel electrophoresis, buffer (50 mM Tris–HCl, pH 8.8, 6 M urea), 30% glycerol, 65 mM dithiothreitol reducing solution, alkylating solution of iodoacetamide at 100 mM, and SDS-PAGE gel at 12% acrylamide were used. The polyacrylamide gel was prepared in the presence of TEMED, a polymerization agent, and ammonium persulfate. Sodium dodecyl sulfate at 2% was used to denature proteins. Migration was carried out under the action of a constant electric field of 25 mA for 15 min followed by 30 mA for ≈5 h. Silver nitrate was used for protein staining. Proteins of interest were recovered by cutting the gel.

Mass Spectrometry

For global proteomic analysis, the protein-containing solution was subjected to dialysis and trypsin digestion. Dialysis was carried out using Slide-ALyzer 2K MWCO dialysis cassettes (Pierce Biotechnology, Rockford, IL, United States) against a solution of 1 M urea and 50 mM ammonium bicarbonate pH 7.4, twice, during 4 h, and one night. Protein digestion was carried out by adding 2 μg of trypsin solution (Promega, Charbonnières, France) to the alkylated proteins, with incubation at 37°C overnight in a water bath. The digested sample was then desalted using detergent columns (Thermo Fisher Scientific, Illkirch, France) and analyzed by mass spectrometry on a Synapt G2Si Q-TOF traveling wave mobility spectrometer (Waters, Guyancourt, France) as described previously (Reteno et al., 2015). An internal protein sequence database was used that was built primarily with two types of amino acid sequences: (i) sequences obtained by translating *P. massiliensis* open reading frames (ORFs); (ii) sequences obtained by translating the whole genome into the six reading frames then fragmenting the six translation products into 250 amino acid-long sequences with a sliding step of 30 amino acids. Contiguous sequences positive for peptide detection were fused and re-analyzed.

Genome Annotation

Gene predictions were performed using GeneMarkS and Prodigal softwares, and results were merged (Besemer and Borodovsky, 2005; Hyatt et al., 2010). ORFs shorter than 50 amino acids were discarded. Predicted proteins were annotated by comparative genomics by using BLASTp searches against the NCBI GenBank non-redundant protein sequence database (nr), with an e-value threshold of 1e−3. ORFans were defined as ORFs without homolog in the nr database considering as thresholds an e-value of 1e−3 and a coverage of the query sequences by alignments of 30%. Functional annotation was refined by using DeltaBLAST searches (Boratyn et al., 2012). Best reciprocal hits were detected by the Proteinortho program with an amino acid identity percentage and a coverage thresholds of 30 and 70%, respectively (Lechner et al., 2011). The core genome and the pangenome were estimated by clustering predicted proteins with CD-HIT (Huang et al., 2010) using 30 and 50% as thresholds for sequence identity and coverage, respectively. Transfer RNAs (tRNAs) were predicted using Aragorn (Laslett and Canback, 2004).

Transcriptomic Analysis for *Pandoravirus massiliensis*

Reads generated from the RNA extracts were mapped on the assembled genome by using the bowtie2 software with default parameters (Langmead et al., 2009; Langmead, 2010; Langmead and Salzberg, 2012). Mapping results were analyzed using the HTseq-count software, with the union mode (Anders et al., 2015). Only "aligned" results were taken into account. Predicted ORFs were considered as transcribed if at least 10 reads were aligned.

Search for Transposable Elements

Miniature inverted repeat transposable elements (MITE) previously identified in the *P. salinus* genome were searched for by using the BLASTn program with an evalue threshold of 1e−3 (Sun et al., 2015). MITE are DNA transposons whose size ranges between 100 and 600 bp and that require transposition enzymes from other, autonomous transposable elements.

Phylogenetic Analyses and Hierarchical Clustering

Phylogeny reconstruction was performed based on the DNA-dependent RNA polymerase subunit 1. Amino acid sequences were aligned using Muscle (Edgar, 2004). The phylogenetic tree was built using FastTree with default parameters (Price

et al., 2010). Hierarchical clustering was performed with the Mev program (Chu et al., 2008) based on the presence/absence patterns of pandoravirus genes that are homologous to clusters of orthologous groups of proteins previously delineated for nucleocytoplasmic large DNA viruses and giant viruses of amoebae (NCVOGs) (Yutin et al., 2013).

RESULTS

Three new pandoravirus isolates were obtained from soil and water samples collected in Brazil in 2015–2016. Two pandoraviruses were isolated in 2015 from soil samples collected from Pampulha lagoon and Belo Horizonte city. A third pandoravirus was isolated in 2016 from a Soda lake (Soda lake2). These new viruses were named *Pandoravirus massiliensis* strain BZ81 c (**Figure 1a**), *Pandoravirus pampulha* strain 8.5 (**Figure 1b**), and *Pandoravirus braziliensis* strain SL2 (**Figure 1c**), respectively.

For the *P. massiliensis* genome, 403,592 reads were obtained by the mate-pair sequencing, with a length ranging from 35 to 251 nucleotides, and the average quality per read was 28 and 37 for the forward and the reverse sequences, respectively. For the paired-end sequencing, 269,656 reads were obtained with a length ranging from 35 to 251 nucleotides; the average quality score per read was 37 for the forward and the reverse sequences, respectively. The *P. massiliensis* genome (EMBL Accession no. OFAI01000000) was assembled in two scaffolds of 1,593,057 and 2,489 bp, and was predicted to encode 1,414 proteins (**Table 1**). Mean size (±SD) of these proteins is 299 ± 228 amino acids. Median size is 218 amino acids. A total of 25% of these predicted proteins are smaller than 136 amino acids, and 25% are larger than 397 amino acids, among which 15 proteins are larger than 1,000 amino acids. Among these 1,414 proteins, 786 (56%) have a homolog in the NCBI GenBank nr database (using a BLASTp e-value threshold of 1e−3), and 628 (44%) are ORFans (ORFs with no significant homolog in the NCBI nr database). Among ORFs that have a homolog in nr, 744 (95%) have genes from previously described pandoraviruses as best BLASTp hits. Two genes encode for Pro-tRNA and Cys-tRNA. A total of 74 ORFs have a significant BLASTp hit with a NCVOG. A total of 310 ORFs were found to be paralogous genes. Finally, 425 ORFs (30% of the gene content) belong to the strict core genome delineated for the six pandoraviruses. For the *P. pampulha* genome, a total of 864,982 reads were obtained, with a length ranging from 35 to 251 nucleotides; the average quality score per read was 37 for the forward and the reverse strands, respectively. The *P. pampulha* genome (EMBL Accession no. OFAJ01000000) was assembled in a single scaffold of 1,676,092 bp, and predicted to encode 2,368 proteins and two tRNA, a Pro-tRNA, and a Trp-tRNA (**Table 1**). Mean size of these proteins is 237 ± 219 amino acids. Among these ORFs, 58% have no homolog in the nr database. Among the 989 ORFs that have a homolog in nr, 974 (98%) have genes from previously described pandoraviruses as best BLASTp hits. A total of 72 ORFs have a hit with a NCVOG. We detected that 407 ORFs (17%) are paralogs. Finally, 417 ORFs (18%) were found to belong to the strict core genome of the pandoraviruses. For the *P. braziliensis* genome, a total of 542,496 reads with a length ranging from 35 to 251 nucleotides were obtained for the paired-end run; the average quality score per read was 37 for the forward and the reverse sequences, respectively. For the mate-pair run, a total of 2,194,091 reads were obtained, with a length ranging from 35 to 251 nucleotides; the average quality score per read was 37 for the forward and the reverse strands, respectively. The assembly of the *P. braziliensis* genome (EMBL Accession no. OFAK01000000) provided two scaffolds with a length of 1,828,953 and 21,873 bp (**Table 1**). A total of 2,693 proteins were predicted, their mean size being 215 ± 212 amino acids. Three genes encode a Leu-tRNA, a Pro-tRNA, and a Pyl-tRNA. ORFans represent 67% of the ORF set. Among the 892 ORFs that have a homolog in nr, 872 (98%) have genes from previously described pandoraviruses as best BLASTp hits. Moreover, 72 ORFs are homologous to a NCVOG. We detected that 437 ORFs are paralogs. Finally, 428 ORFs (16%) were found to be shared with the five other pandoraviruses. All these three genomes were found to be linear double-stranded DNA, as described previously for *P. salinus*, *P. dulcis*, and *P. inopinatum*. Thus, here, PCR amplification performed with the attempt to test the circularity of the genome failed. BLASTp hits were found in nr for hundreds of additional short ORFs predicted in the genomes of *P. pampulha* and *P. braziliensis*, but e-values were >1e−3, and only short fragments from these sequences were usually involved in alignments obtained with these hits.

The pangenome size delineated for these three new pandoravirus genomes and the three previously described pandoravirus genomes reaches 7,477 gene comprising clusters or unique genes (**Figures 2**, **3**). Among them, 6,108 (82%) encompass a single predicted gene (**Figure 4**). A total of 427 clusters (5.7%) are composed of two representative sequences and 163 clusters (2.2%) are composed of three representative sequences. The "strict" core genome represents 4.7% of the pangenome. It includes 352 clusters comprising 2,617 pandoravirus proteins, each of these clusters encompassing at least one predicted protein from each of the six pandoravirus isolates. The ratio core genome/pangenome is thus less than 0.05 and the proportion for each individual virus of the gene content that belongs to the core genome is comprised between 15.4 and 29.4%. When considering the proteins involved in best reciprocal hits with an identity >30% and a query sequence coverage >70%, a total of 208 clusters of proteins (1.6% of the full cluster set) encompassed at least one protein of each of the six pandoravirus isolates. Besides, a homolog was found in the gene content of all six pandoraviruses for a NCVOG in 403 cases.

Only 13% of the *P. massiliensis* transcripts were detected during the first 4 h post-infection of the amoeba by this virus. In contrast, more than two-thirds of the transcripts (69%) were detected 6 h post-infection of the amoeba, and 18% of them were detected 8 h post-infection. A total of 359 *P. massiliensis* ORFs (25% of the gene content) were detected by transcriptomics taking into account all reads at any time post-infection, with a mean coverage of 50 reads/ORF along the whole genome. Among these 359 ORFs, three (ORFs 1, 1,350, and 1,364) had a particularly high coverage, greater than 1,200 reads/ORF (1,592, 1,243, and 1,234, respectively). When removing these three ORFs,

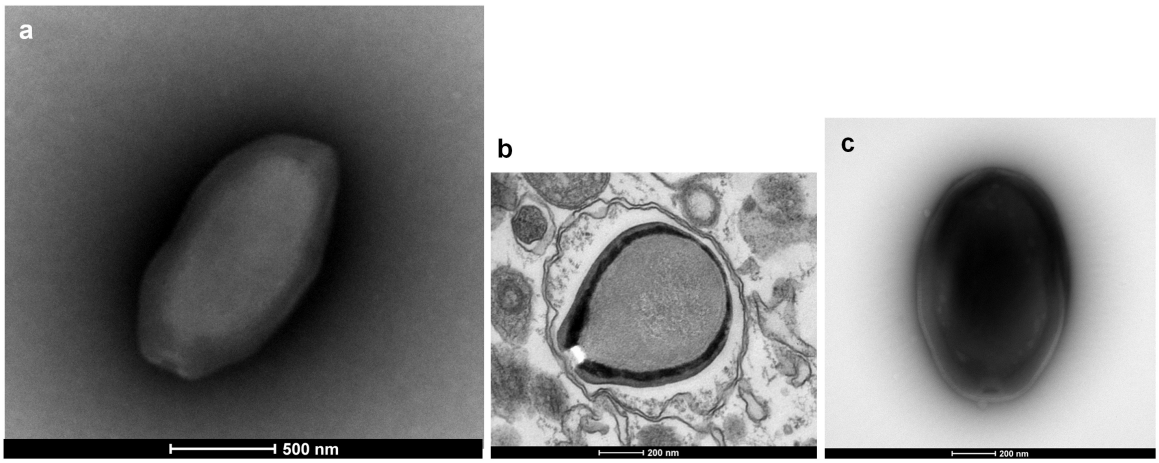


FIGURE 1 | Electron microscopy pictures of pandoravirus isolates by negative staining **(a,c)** or after inclusion **(b)**. **(a)** *Pandoravirus massiliensis*; **(b)** *Pandoravirus pampulha*; **(c)** *Pandoravirus braziliensis*.

TABLE 1 | Main features of the six pandoravirus genomes.

Virus	Genome length	Number of scaffolds	Number of predicted proteins	GC%	Proportion of the gene content that belongs to the core genome (%) (number of proteins)
Pandoravirus dulcis	1,908,520	1	1,487	63.7	29.4 (437)
Pandoravirus salinus	2,473,870	1	2,541	61.7	17.3 (440)
Pandoravirus inopinatum	2,243,110	1	1,839	60.7	26.9 (495)
Pandoravirus pampulha	1,676,092	1	2,368	63.9	17.6 (416)
Pandoravirus braziliensis	1,850,826	2	2,693	59.0	15.4 (415)
Pandoravirus massiliensis	1,595,546	2	1,414	60.1	29.3 (414)

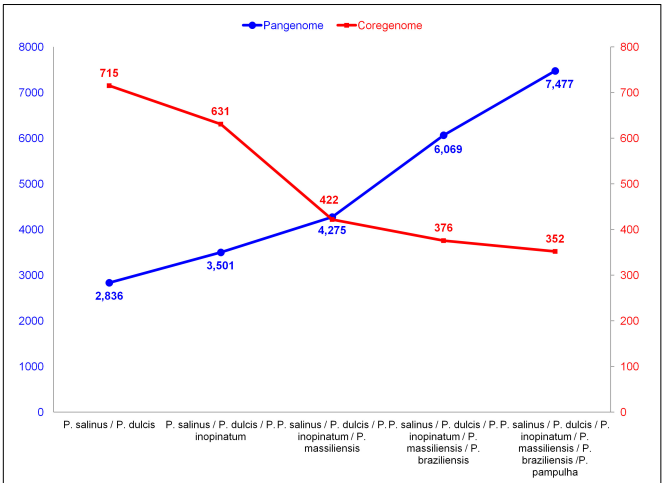


FIGURE 2 | Evolution of the size of the pangenome and core genome of pandoraviruses with the description of each of the six pandoravirus isolates.

the mean coverage of transcripts along the genome decreased to 39 reads/ORF. Two of these three ORFs are hypothetical proteins and were detected in the five other pandoraviruses. Nevertheless, the product of only one of these two ORFs was found by proteomics. Strikingly, this ORF is harbored by the 2,489 bp-long genomic fragment. The second of these two ORFs is contiguous to two other highly transcribed genes (with 425 and 552 mapped reads). The third most transcribed ORF is a collagen triple helix encoding protein, also found by proteomics. Finally, a total of 210 of the 359 transcribed ORFs (58%) is part of the core genome; while 60% of the ORFs that are part of the core genome were transcribed. Conversely, only 149 (14%) of the 1,062 *P. massiliensis* ORFs that do not belong to the core genome were transcribed.

A total of 162 ORFs were found by proteomic analysis of the *P. massiliensis* virions. Among them, 90 proteins (55%) are part of the core genome. Conversely, a protein was found in *P. massiliensis* virions for only 72 (7%) of the 1,062 ORFs that did not belong to the core genome. In addition, the products of 28 ORFans and 99 hypothetical proteins were part from these 162 proteins detected by proteomic analyses. The most abundant peptides found in the *P. massiliensis* virions match with 37 proteins, which include 12 ORFan gene products; 19 hypothetical proteins; a trimeric LpxA-like enzyme motif-containing protein; a translation initiation inhibitor belonging to the YJGF family; a thioredoxin-like fold motif-containing protein; a laminin G domain-containing protein; a collagen triple helix repeat domain-containing protein; and an ankyrin repeat-containing protein. A concordance

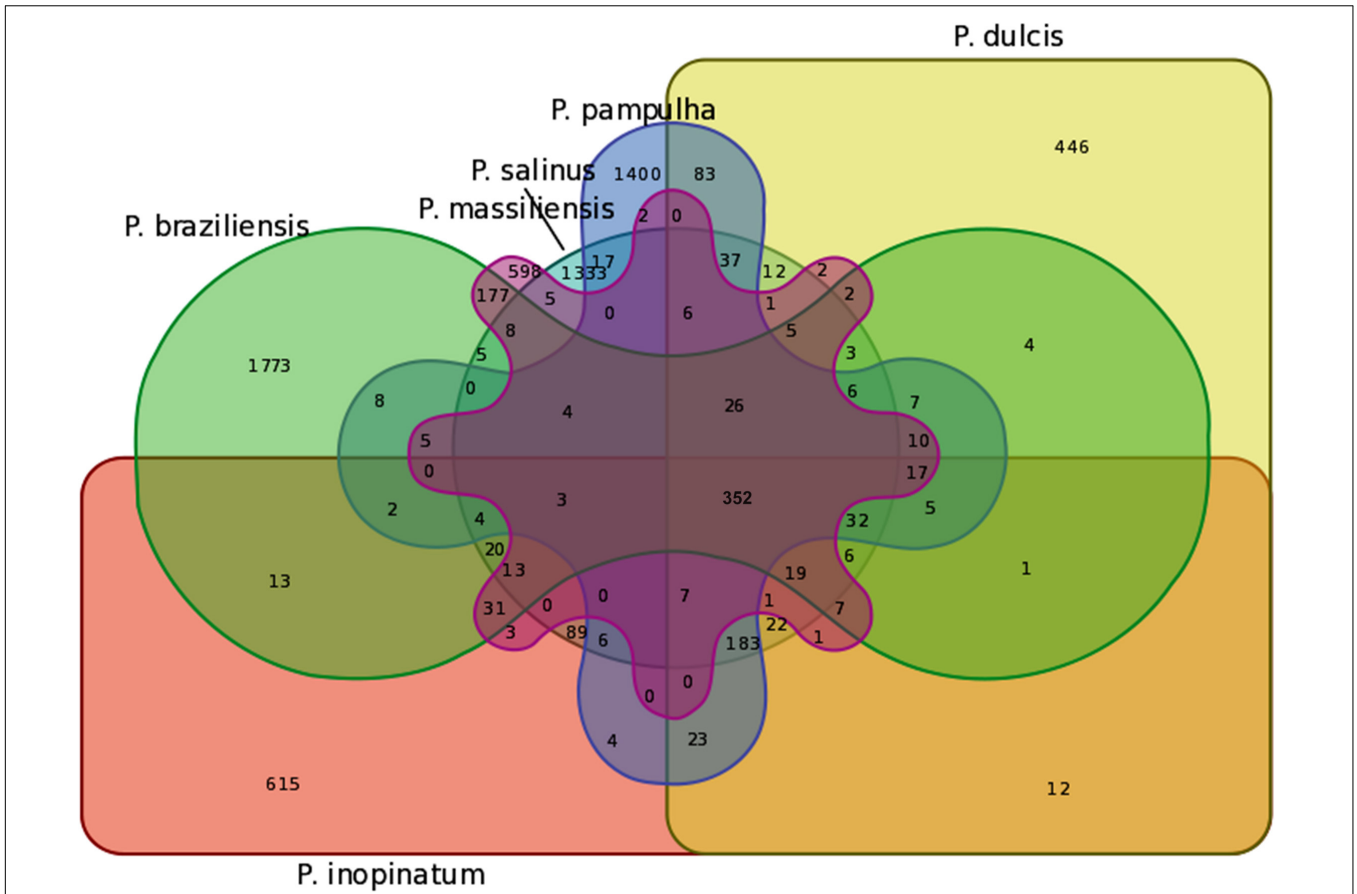


FIGURE 3 | Venn diagram of genes shared and not shared between the gene contents of the six pandoravirus isolates. Venn diagram was built using the following online tool: http://bioinformatics.psb.ugent.be/webtools/Venn/.

between transcriptomic and proteomic data was found for 89 ORFs (**Supplementary Table S1**). These ORFs include 2 ORFans and 61 hypothetical proteins, all found in other pandoraviruses. The other ORFs with functional annotations have a pandoravirus protein as their most similar sequence. These ORFs notably encode an acid phosphatase class b; a C1q domain-containing protein; two casein kinases; a cathepsin c1-like peptidase; a trypsin-like serine protease; a disulfide isomerase motif-containing protein; a DNA pol III gamma/tau subunit-like domain containing protein; an FAD/FMN-containing dehydrogenase; a hexapeptide repeat-containing protein; a histidine phosphatase motif-containing protein; a laminin G domain-containing protein; a lipase/esterase; an NAD-dependent amine oxidase; an oxidoreductase; an SMC ATPase domain-containing protein, SMC proteins being ATPases involved in chromosome organization and dynamics; a thioredoxin-like fold motif-containing protein; two translation initiation inhibitors belonging to the YJGF family; and a trimeric LpxA-like enzyme motif-containing protein (bacterial transferase). Of note, for the five genes predicted to encode DNA-dependent RNA polymerase subunits, transcripts were only detected for those encoding subunits 1 and 2 and no protein was detected by proteomics. Finally, among *P. massiliensis* ORFans, 26 (4.1%) were found to be transcribed, and a similar number (28; 4.5%) were found to encode proteins detected in virions.

Sequences similar to MITEs were identified through BLAST searches, in all six genomes of pandoraviruses, albeit their number varied considerably according to the genome. Thus, eight different matches with MITEs were identified in the *P. massiliensis* genome, which displayed a nucleotide identity varying between 78 and 100% with a MITE identified in *P. salinus* (Sun et al., 2015). Seven matches with MITEs were detected in the *P. inopinatum* genome, which were 76–100% identical with a *P. salinus* MITE. Five matches with MITEs were detected in *P. braziliensis* and *P. dulcis*, which were 75–98 and 76–95% identical with a MITE identified in *P. salinus*, respectively. Finally, four matches with MITEs were identified in the *P. pampulha* genome, which were 82–98% identical with a *P. salinus* MITE. However, when considering full-length MITE copies described for the *P. salinus* genome, 3 and 2 such full-length MITEs were detected in the genomes of *P. massiliensis* and *P. inopinatum*, respectively (**Figure 5**). These sequences did not cluster together according to the isolate (**Supplementary Figure S1**).

Phylogenetic reconstruction based on the RNA polymerase subunit 1 showed that *P. massiliensis* and *P. braziliensis*

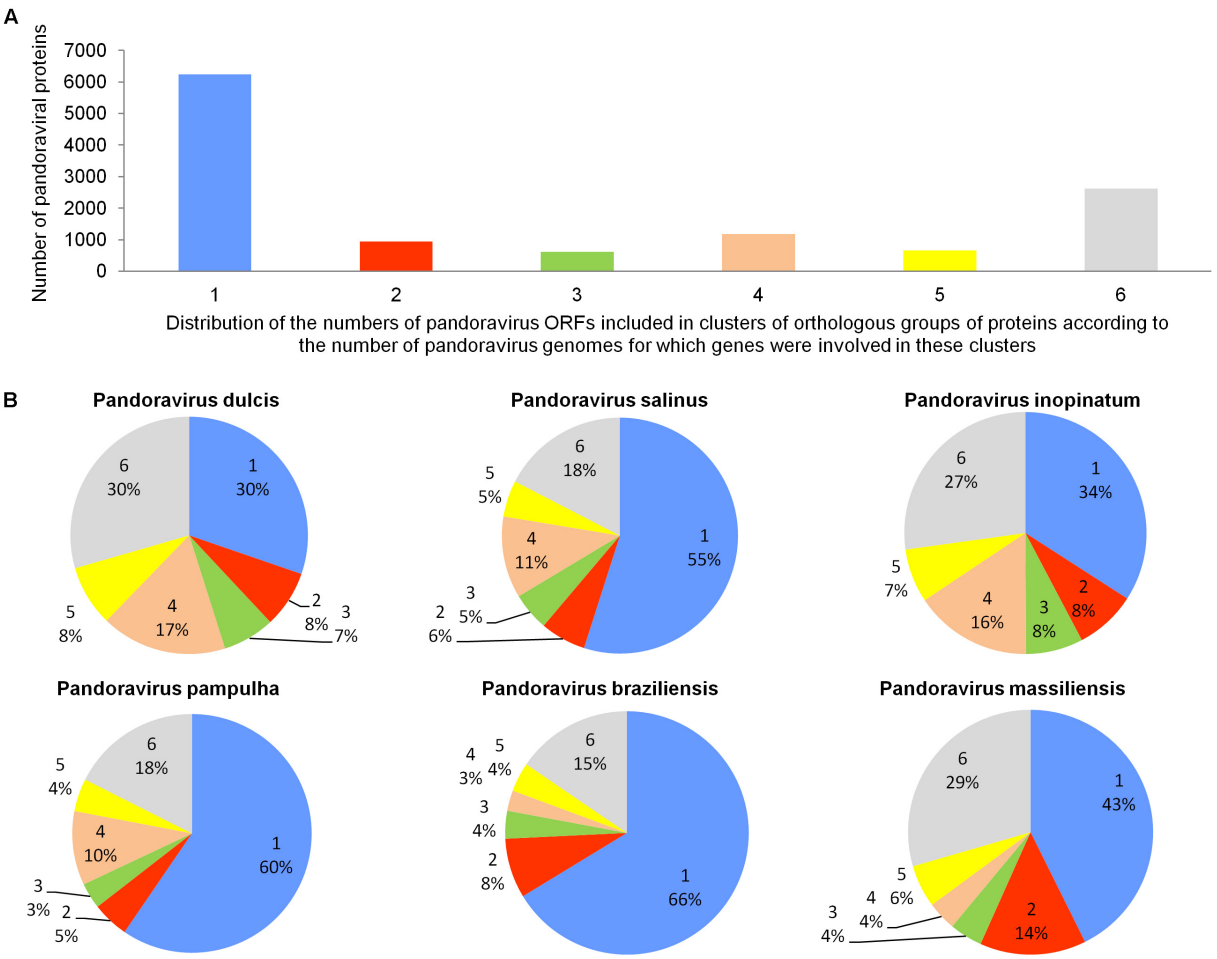


FIGURE 4 | Distribution of the number of pandoravirus ORFs included in clusters of orthologous groups of proteins according to the number of pandoravirus genomes for which genes were involved in these clusters **(A)** and proportion for each pandoravirus genome of the genes involved in clusters including genes from one to six pandoraviruses **(B)**.

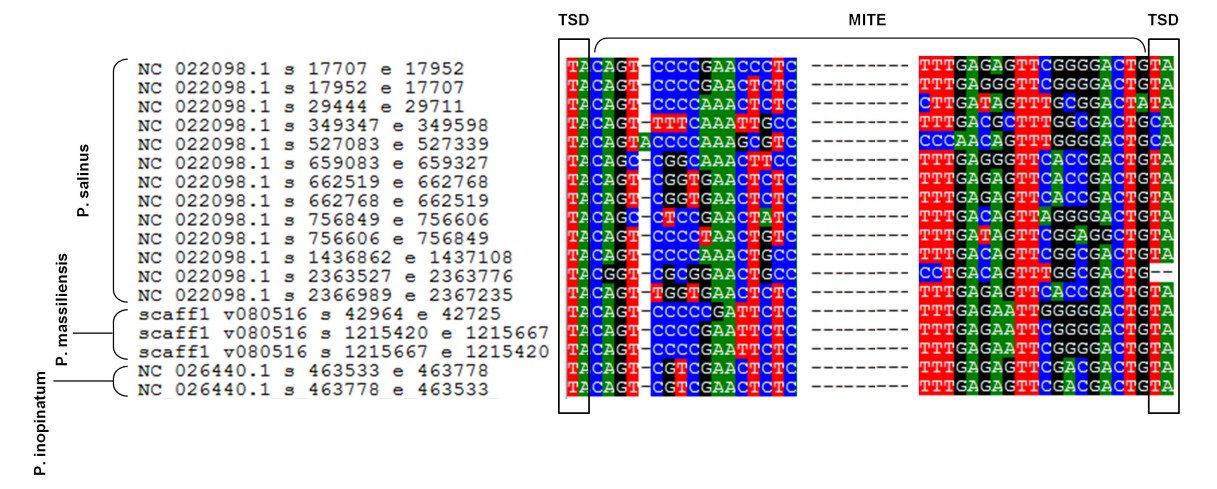


FIGURE 5 | Alignment by muscle of the 13 full-length miniature inverted repeat transposable elements (MITEs) described in the genome of *Pandoravirus salinus* (Sun et al., 2015) and those detected in the genomes of *P. massiliensis* and *P. inopinatum*. Alignment is limited to regions including the flanking target site duplication (TSD) sequences and the start and end of the MITEs.

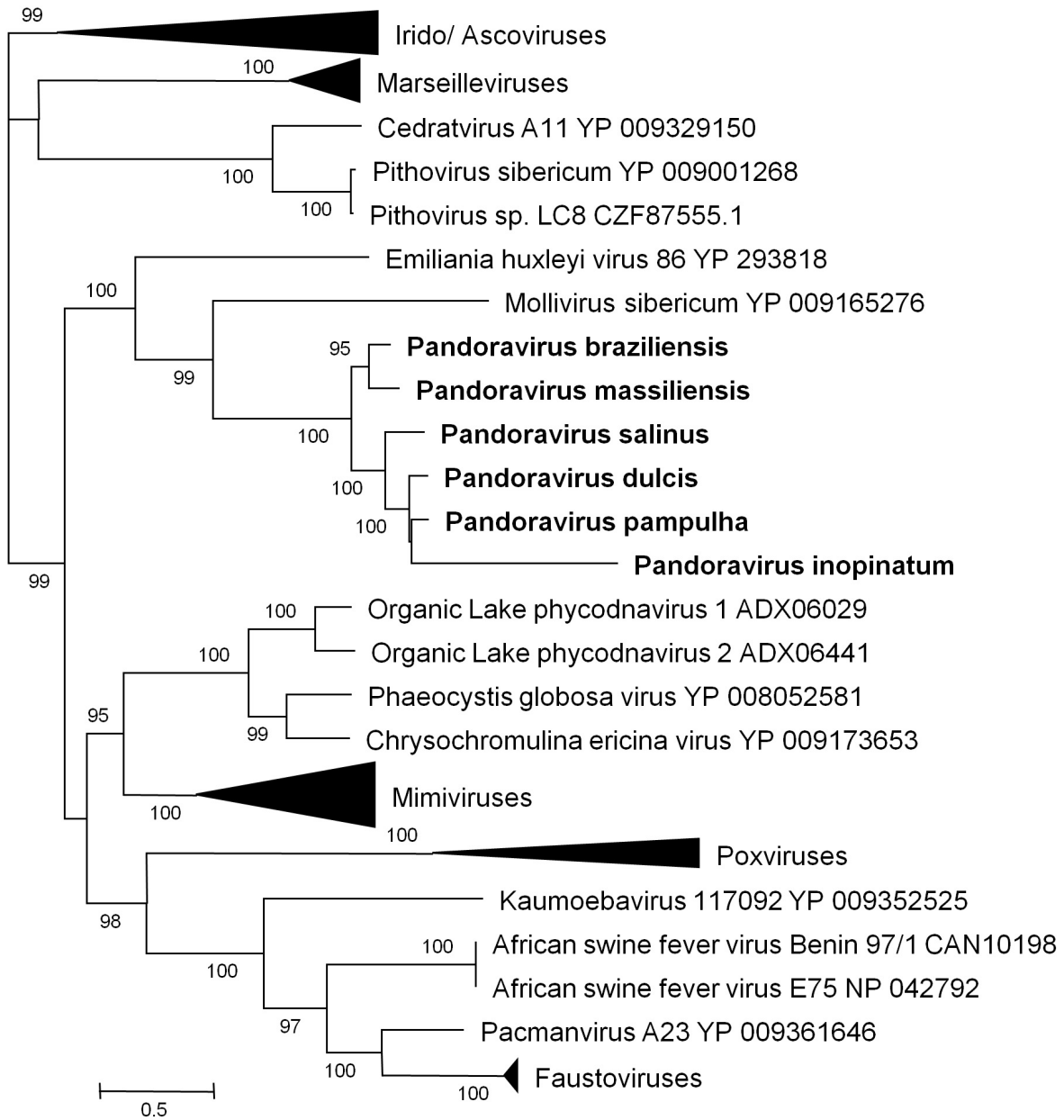


FIGURE 6 | Phylogenetic reconstruction based on amino acid sequences of the DNA-dependent RNA polymerase subunit 1 from representatives of megavirales. Phylogenetic tree was drawn using the maximum likelihood model with the FastTree program (Price et al., 2010).

were closely related (**Figure 6**). Hierarchical clustering showed congruent results with a close relationship between *P. massiliensis* and *P. braziliensis* (**Figure 7**). In addition, mean amino acid identities between orthologous proteins of *P. salinus* and *P. massiliensis* or *P. braziliensis* were similar (mean values, 50.0 and 51.1%, respectively), and lower than the mean amino acid identity between orthologous proteins of *P. salinus* and *P. pampulha* (61.0%) (**Figure 8**). Taken together, on the basis of phylogenetic analysis, the presence/absence patterns of clusters of orthologous groups of proteins of Megavirales members, and amino acid identity of orthologous proteins, two major groups can be delineated for these six pandoravirus isolates. The first group is comprised by *P. massiliensis* and braziliensis, and the second group is comprised by *P. salinus*, *P. dulcis*, *P. pampulha*, and *P. inopinatum*.

Comparison of genome architecture and co-linearity showed a general tendency among the different pandoravirus genomes for a greater co-linearity around the first third of the genome alignement by the MAUVE software, displaying large blocks with a high level of nucleotide identity (**Figure 9**). Besides, dot plots constructed separately for the three new pandoravirus isolates described here on the basis of their gene content showed a

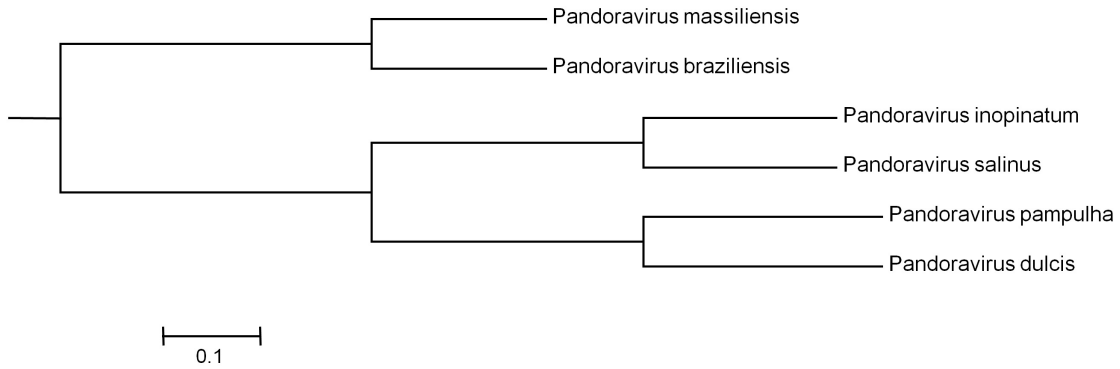


FIGURE 7 | Hierarchical clustering based on the presence/absence patterns of clusters of orthologous groups of proteins of megavirales members in the genomes of pandoraviruses.

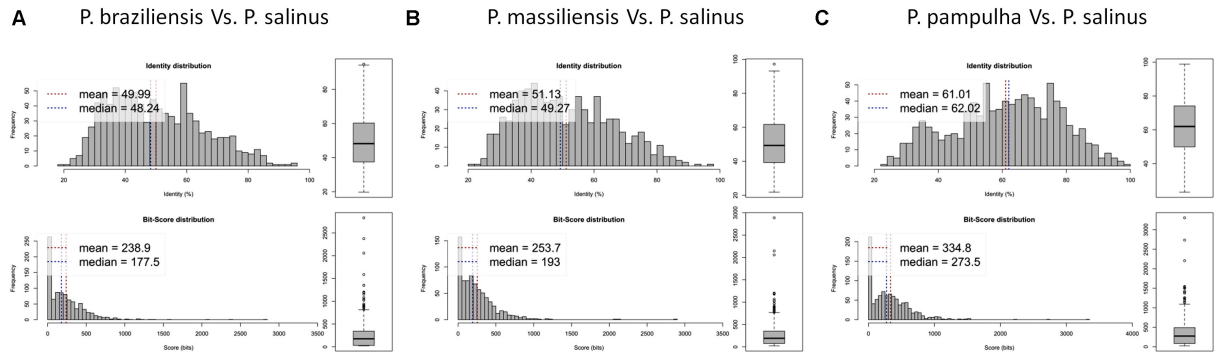


FIGURE 8 | Average amino acid identity between ORFs predicted for pandoravirus genomes. For each comparison, estimates were obtained using both best hits and reciprocal best hits between two sets of proteins from a pandoravirus newly described here and *Pandoravirus salinus*. **(A)** *Pandoravirus braziliensis* versus *Pandoravirus salinus*, **(B)** *Pandoravirus massiliensis* versus *Pandoravirus salinus*, **(C)** *Pandoravirus pampulha* versus *Pandoravirus salinus*.

considerable number of paralogous genes, and the scattering of core genes along the whole genome length (**Figure 10**). Paralogous genes mostly consisted in three groups of proteins with ankyrin repeat motifs, F-box domains, and MORN-repeats. Finally, the gene of *P. salinus* recently described as a putative candidate for encoding a capsid protein (ps_862) (Sinclair et al., 2017) was detected in the genomes of *P. braziliensis*, *P. massiliensis*, and *P. pampulha*. However, the product of this gene was not found in the proteome of *P. massiliensis* virions.

DISCUSSION

We delineated here the pangenome and core genome of pandoraviruses based on six viruses, including three new isolates from Brazil. Our findings indicate that pandoraviruses, first described in 2013, are likely common in water and soil samples worldwide, as is the case for mimiviruses and marseilleviruses. The various pandoravirus isolates described to date were isolated from three continents in Chile, Australia, Germany, and Brazil (Philippe et al., 2013; Scheid et al., 2014; Dornas et al., 2016; Andrade et al., 2018). Moreover, our results indicate that pandoraviruses currently form a homogenous viral group, regarding both their morphology and their genome organization and content.

Our findings further point out that these giant viruses are currently those with the largest genomes, which range in size from 1.59 Mbp (for *P. massiliensis*) to 2.47 Mbp (for *P. salinus*). Far smaller genomes have been described for other giant viruses, namely pithoviruses (Legendre et al., 2014; Levasseur et al., 2016) and cedratviruses (Andreani et al., 2017; Bertelli et al., 2017). Indeed, genome size is 0.61–0.68 Mbp for pithoviruses and 0.57–0.59 Mbp for cedratviruses. This is intriguing as the size of pithovirus and cedratvirus virions, which have a similar morphology than pandoravirus virions and a similar tegument-resembling structure delineating the particle, is similar to those of pandoravirus virions, or even larger for pithoviruses (up to 1.5–2.5 μm compared to c.a. 1 μm for pandoraviruses) (Legendre et al., 2015; Okamoto et al., 2017). Such discrepancies between genome and virion sizes have been rarely described (Cui et al., 2014; Brandes and Linial, 2016).

We noted here a great size of the pandoravirus pangenome (comprised by 7,477 unique genes or clusters of genes), compared with that delineated most recently for mimiviruses (2,869 clusters) (Assis et al., 2017) and marseilleviruses (665 clusters) (Dornas et al., 2016). Furthermore, expansion of this

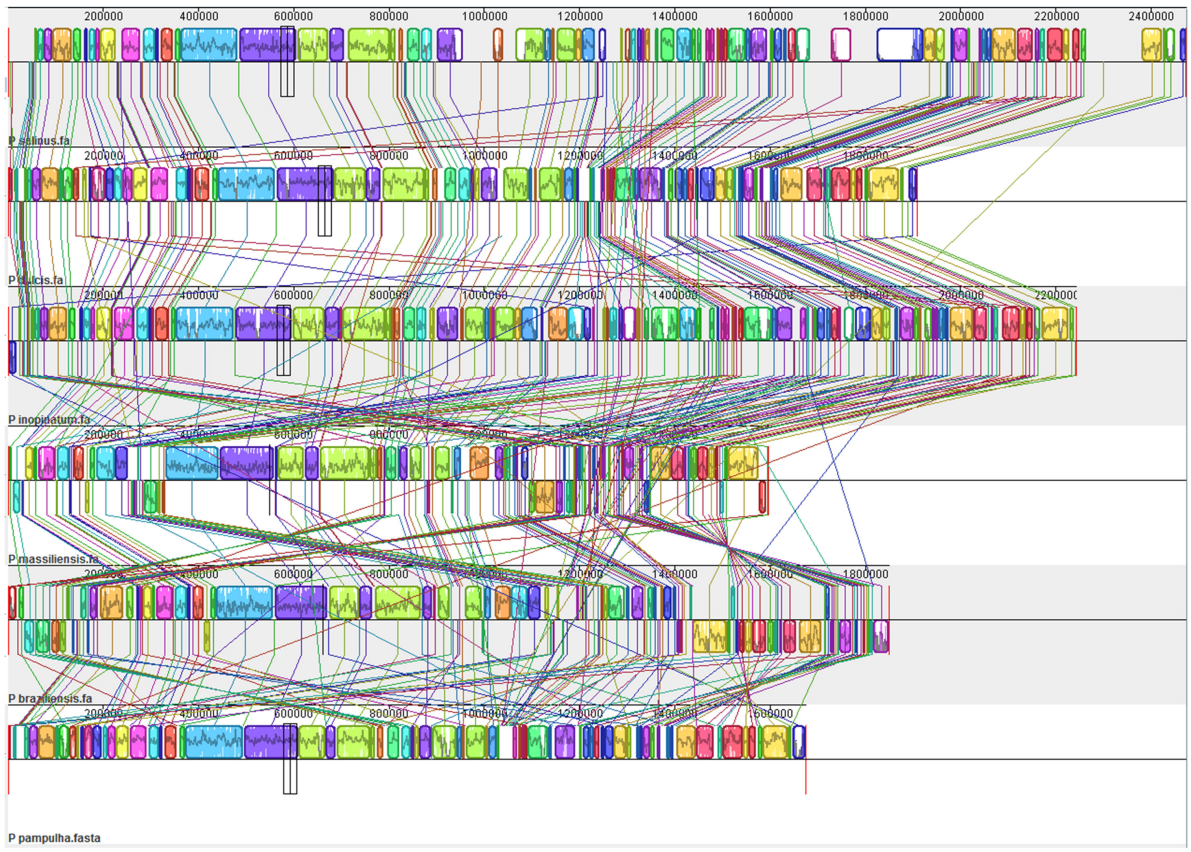


FIGURE 9 | Whole genome alignment of pandoravirus genomes by the MAUVE program (Darling et al., 2010).

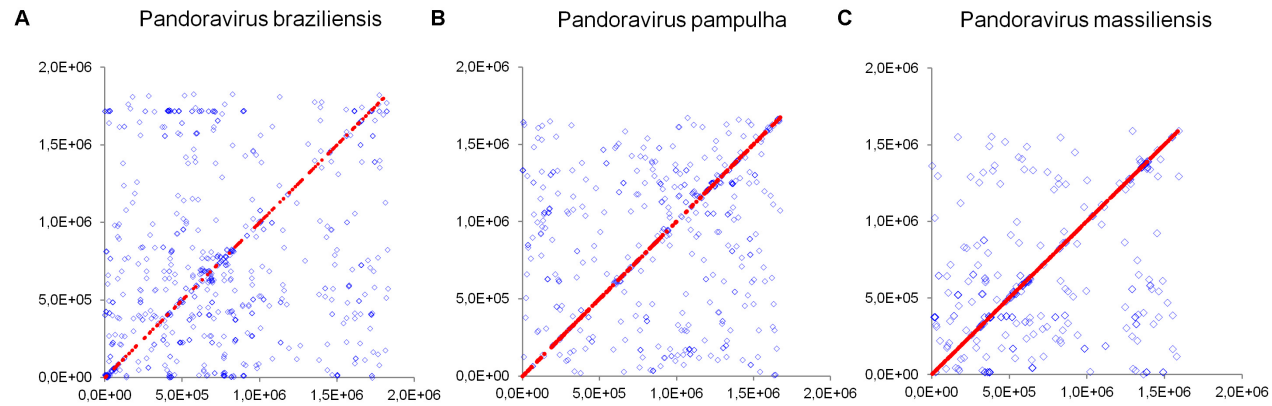


FIGURE 10 | Distribution of core genes and paralogous genes along the pandoravirus genomes. **(A)** *Pandoravirus braziliensis*; **(B)** *Pandoravirus pampulha*; **(C)** *Pandoravirus massiliensis*. Core genes are indicated by red dots; paralogous genes are indicated by blue diamonds.

pangenome since 2013, while taking into account the three new pandoravirus genomes described here, suggests it is still open with a mean increase of 28% at each new genome annotation. Conversely, a major finding of our pangenome analysis is that pandoraviruses have a core genome size that is limited relatively to the number of genes predicted in each of their genomes. Thus, the proportion for each individual virus of the gene content that belongs to the core genome is lower than 30% and as low as 15%. Compared to the 352 clusters of genes described for the pandoravirus core genome, mimiviruses core genome comprises 267 clusters of genes based on 21 described genomes with a size ranging between 1,017 and 1,259 Mbp (Assis et al., 2017) and the marseillevirus core genome comprises 202 clusters of genes based on 8 described genomes

with a size ranging between 0.347 and 0.386 Mbp (Dornas et al., 2016).

Strikingly, a significant number of pandoravirus predicted ORFs have no homolog in the international databases and no predicted functions. This proportion of ORFans remains greater than for other giant viruses of amoebae (Colson et al., 2017b). The *P. massiliensis* transcriptomic and proteomic analyses showed that at least a small proportion of these ORFan genes are transcribed and encode for proteins. This highlights that most of the gene armentarium involved in the structure, metabolism, and replication of these pandoraviruses is currently unknown, as is the case for all other giant viruses of amoebae. We also noted that coding capacity differed greatly from one pandoravirus genome to another. Thus, *P. braziliensis* harbors the biggest gene content with a total of 2,693 predicted genes and a coding capacity of 1.45 gene/kbp. In contrast, *P. dulcis*, with a genome of similar size, is predicted to encode only 1,502 genes, corresponding to a coding capacity of 0.79 gene/kbp. Regarding the genomes of the three new pandoraviruses, the mean size of their genes varies greatly, from 215 to 299 amino acids. Moreover, the gene contents of the three new pandoraviruses differ in terms of proportions of ORFans, ranging betweeen 44 and 67%.

The presence of MITEs in the pandoravirus genomes are another evidence of the presence of transposable elements in the genomes of giant viruses of amoebae. Previously, transpovirons were described in mimivirus genomes, and genomes of virophages were found to integrate as provirophages in the genomes of these mimiviruses (Desnues et al., 2012). Moreover, introns were described in genomes of several giant viruses of amoebae (Desnues et al., 2012; Philippe et al., 2013; Colson et al., 2017b). Taken together, all these elements correspond to a mobilome for these giant viruses (Desnues et al., 2012). In addition to full-length MITEs, we detected several sequences in the different pandoravirus genomes that match with full-length MITEs. They might correspond to degraded MITE sequences or to different elements. Besides, two ribonuclease H-like domain motif-containing proteins were detected as part of the transcriptome of *P. massiliensis*. This deserves being mentioned since the presence of ribonuclease H in the genomes of giant virus has been recently studied and suggested to be associated with sequence integration (Moelling et al., 2017).

In summary, our knowledge of the pandoravirus diversity continues to expand (Andrade et al., 2018). Further analyses should allow to gain a better knowledge and understanding of the evolution and origin of these giant pandoraviruses, and of their relationships with viruses and cellular microorganisms.

AUTHOR CONTRIBUTIONS

PC, BS, DR, and SA designed the experiments. SA, PC, JAN, EB, AO, FD, AA, and AL contributed to the data and performed the experiments. SA, PC, JAB, EC, AL, DR, and BS analyzed the data. PC, BS, and SA wrote the manuscript.

FUNDING

This work was 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), and by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

ACKNOWLEDGMENTS

We are thankful to Hiroyuki Hikida, Flora Marchandise, Saïd Azza, Philippe Decloquement, and Caroline Blanc-Tailleur for their technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01486/full#supplementary-material

FIGURE S1 | Molecular phylogenetic analysis of miniature inverted repeat transposable elements (MITEs) detected in the genomes of *P. salinus*, *P. massiliensis* and *P. inopinatum*. The tree was built with nucleotide sequences from **Figure 5**, using the Maximum Likelihood method. Blue squares indicate sequences of *P. salinus*; green circles indicate sequences of *P. massiliensis*; gray losanges indicate sequences of *P. inopinatum*.

TABLE S1 | *Pandoravirus massiliensis* predicted genes for which a transcript has been detected by transcriptomics and a product has been detected by proteomic.

REFERENCES

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Andrade, A. C. D. S. P., Arantes, T. S., Rodrigues, R. A. L., Machado, T. B., Dornas, F. P., Landell, M. F., et al. (2018). Ubiquitous giants: a plethora of giant viruses found in Brazil and Antarctica. *Virol. J.* 15:22. doi: 10.1186/s12985-018-0930-x

Andreani, J., Aherfi, S., Bou Khalil, J. Y., Di Pinto, F., Bitam, I., Raoult, D., et al. (2016). Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. *Viruses* 8:E300. doi: 10.3390/v8110300

Andreani, J., Khalil, J. Y. B., Sevvana, M., Benamar, S., Di Pinto, F., Bitam, I., et al. (2017). Pacmanvirus, a new giant icosahedral virus at the crossroads between *Asfarviridae* and faustoviruses. *J. Virol.* 91:e00212-17. doi: 10.1128/JVI.00212-17

Assis, F. L., Franco-Luiz, A. P. M., Dos Santos, R. N., Campos, F. S., Dornas, F. P., Borato, P. V. M., et al. (2017). Genome characterization of the first mimiviruses of lineage C isolated in Brazil. *Front. Microbiol.* 8:2562. doi: 10.3389/fmicb.2017.02562

Atoni, E., Wang, Y., Karungu, S., Waruhiu, C., Zohaib, A., Obanda, V., et al. (2018). Metagenomic virome analysis of culex mosquitoes from Kenya and China. *Viruses* 10:E30. doi: 10.3390/v10010030

Bertelli, C., Mueller, L., Thomas, V., Pillonel, T., Jacquier, N., and Greub, G. (2017). Cedratvirus lausannensis - digging into Pithoviridae

diversity. *Environ. Microbiol.* 19, 4022–4034. doi: 10.1111/1462-2920.13813

Besemer, J., and Borodovsky, M. (2005). GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* 33, W451–W454. doi: 10.1093/nar/gki487

Boratyn, G. M., Schaffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., and Madden, T. L. (2012). Domain enhanced lookup time accelerated BLAST. *Biol. Direct.* 7:12. doi: 10.1186/1745-6150-7-12

Brandes, N., and Linial, M. (2016). Gene overlapping and size constraints in the viral world. *Biol. Direct.* 11:26. doi: 10.1186/s13062-016-0128-3

Brinkman, N. E., Villegas, E. N., Garland, J. L., and Keely, S. P. (2018). Reducing inherent biases introduced during DNA viral metagenome analyses of municipal wastewater. *PLoS One* 13:e0195350. doi: 10.1371/journal.pone.0195350

Chu, V. T., Gottardo, R., Raftery, A. E., Bumgarner, R. E., and Yeung, K. Y. (2008). MeV+R: using MeV as a graphical user interface for Bioconductor applications in microarray analysis. *Genome Biol.* 9, R118–R119. doi: 10.1186/gb-2008-9-7-r118

Colson, P., La Scola, B., Levasseur, A., Caetano-Anolles, G., and Raoult, D. (2017a). Mimivirus: leading the way in the discovery of giant viruses of amoebae. *Nat. Rev. Microbiol.* 15, 243–254. doi: 10.1038/nrmicro.2016.197

Colson, P., La Scola, B., and Raoult, D. (2017b). Giant viruses of amoebae: a journey through innovative research and paradigm changes. *Annu. Rev. Virol.* 4, 61–85. doi: 10.1146/annurev-virology-101416-041816

Cui, J., Schlub, T. E., and Holmes, E. C. (2014). An allometric relationship between the genome length and virion volume of viruses. *J. Virol.* 88, 6403–6410. doi: 10.1128/JVI.00362-14

Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147

Desnues, C., La Scola, B., Yutin, N., Fournous, G., Robert, C., Azza, S., et al. (2012). Provirophages and transpovirons as the diverse mobilome of giant viruses. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18078–18083. doi: 10.1073/pnas.1208835109

Dornas, F. P., Assis, F. L., Aherfi, S., Arantes, T., Abrahao, J. S., Colson, P., et al. (2016). A Brazilian *Marseillevirus* is the founding member of a lineage in family *Marseilleviridae*. *Viruses* 8:76. doi: 10.3390/v8030076

Dornas, F. P., Khalil, J. Y., Pagnier, I., Raoult, D., Abrahao, J., and La Scola, B. (2015). Isolation of new Brazilian giant viruses from environmental samples using a panel of protozoa. *Front. Microbiol.* 6:1086. doi: 10.3389/fmicb.2015.01086

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26, 680–682. doi: 10.1093/bioinformatics/btq003

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Kerepesi, C., and Grolmusz, V. (2017). The "Giant Virus Finder" discovers an abundance of giant viruses in the Antarctic dry valleys. *Arch. Virol.* 162, 1671–1676. doi: 10.1007/s00705-017-3286-4

Khalil, J. Y., Langlois, T., Andreani, J., Sorraing, J. M., Raoult, D., Camoin, L., et al. (2017). Flow cytometry sorting to separate viable giant viruses from amoeba co-culture supernatants. *Front. Cell. Infect. Microbiol.* 6:202. doi: 10.3389/fcimb.2016.00202

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., et al. (2003). A giant virus in amoebae. *Science* 299:2033. doi: 10.1126/science.1081867

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinformatics* Chapter 11:Unit 11.7. doi: 10.1002/0471250953.bi1107s32

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32, 11–16. doi: 10.1093/nar/gkh152

Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124. doi: 10.1186/1471-2105-12-124

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (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. doi: 10.1073/pnas.1320670111

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (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–E5335. doi: 10.1073/pnas.1510795112

Levasseur, A., Andreani, J., Delerce, J., Bou Khalil, K. J., Robert, C., La Scola, B., et al. (2016). Comparison of a modern and fossil Pithovirus reveals its genetic conservation and evolution. *Genome Biol. Evol.* 8, 2333–2339. doi: 10.1093/gbe/evw153

Moelling, K., Broecker, F., Russo, G., and Sunagawa, S. (2017). RNase H as gene modifier, driver of evolution and antiviral defense. *Front. Microbiol.* 8:1745. doi: 10.3389/fmicb.2017.01745

Okamoto, K., Miyazaki, N., Song, C., Maia, F. R. N. C., Reddy, H. K. N., Abergel, C., et al. (2017). 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.* 7, 13291–13390. doi: 10.1038/s41598-017-13390-4

Philippe, N., Legendre, M., Doutre, G., Coute, Y., Poirot, O., Lescot, M., et al. (2013). Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281–286. doi: 10.1126/science.1239181

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Reteno, D. G., Benamar, S., Khalil, J. B., Andreani, J., Armstrong, N., Klose, T., et al. (2015). Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* 89, 6585–6594. doi: 10.1128/JVI.00115-15

Scheid, P. (2016). A strange endocytobiont revealed as largest virus. *Curr. Opin. Microbiol.* 31, 58–62. doi: 10.1016/j.mib.2016.02.005

Scheid, P., Balczun, C., and Schaub, G. A. (2014). Some secrets are revealed: parasitic keratitis amoebae as vectors of the scarcely described pandoraviruses to humans. *Parasitol. Res.* 113, 3759–3764. doi: 10.1007/s00436-014-4041-3

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19, 1117–1123. doi: 10.1101/gr.089532.108

Sinclair, R. M., Ravantti, J. J., and Bamford, D. H. (2017). Nucleic and amino acid sequences support structure-based viral classification. *J. Virol.* 91:e02275-16. doi: 10.1128/JVI.02275-16

Sun, C., Feschotte, C., Wu, Z., and Mueller, R. L. (2015). DNA transposons have colonized the genome of the giant virus *Pandoravirus salinus*. *BMC Biol.* 13:38. doi: 10.1186/s12915-015-0145-1

Temmam, S., Davoust, B., Chaber, A. L., Lignereux, Y., Michelle, C., Monteil-Bouchard, S., et al. (2017). Screening for viral pathogens in African Simian bushmeat seized at A French airport. *Transbound. Emerg. Dis.* 64, 1159–1167. doi: 10.1111/tbed.12481

Temmam, S., Monteil-Bouchard, S., Sambou, M., Aubadie-Ladrix, M., Azza, S., Decloquement, P., et al. (2015). Faustovirus-like Asfarvirus in hematophagous biting midges and their vertebrate hosts. *Front. Microbiol.* 6:1406. doi: 10.3389/fmicb.2015.01406

Tritt, A., Eisen, J. A., Facciotti, M. T., and Darling, A. E. (2012). An integrated pipeline for de novo assembly of microbial genomes. *PLoS One* 7:e42304. doi: 10.1371/journal.pone.0042304

Verneau, J., Levasseur, A., Raoult, D., La Scola, B., and Colson, P. (2016). MG-digger: an automated pipeline to search for giant virus-related sequences in metagenomes. *Front. Microbiol.* 7:428. doi: 10.3389/fmicb.2016.00428

Yutin, N., Colson, P., Raoult, D., and Koonin, E. V. (2013). Mimiviridae: clusters of orthologous genes, reconstruction of gene repertoire evolution and proposed expansion of the giant virus family. *Virol. J.* 10:106. doi: 10.1186/1743-422X-10-106

Zerbino, D. R. (2010). Using the Velvet de novo assembler for short-read sequencing technologies. *Curr. Protoc. Bioinformatics* Chapter 11:Unit11.5. doi: 10.1002/0471250953.bi1105s31

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

Morphologic and Genomic Analyses of New Isolates Reveal a Second Lineage of Cedratviruses

Rodrigo Araújo Lima Rodrigues,[a,b] Julien Andreani,[a] Ana Cláudia dos Santos Pereira Andrade,[b] Talita Bastos Machado,[b] Souhila Abdi,[a] Anthony Levasseur,[a] Jônatas Santos Abrahão,[b] Bernard La Scola[a]

[a]Aix Marseille Université, IRD, APHM, MEPHI, IHU-Méditerranée Infection, Marseille, France

[b]Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

ABSTRACT Giant viruses have been isolated and characterized in different environments, expanding our knowledge about the biology of these unique microorganisms. In the last 2 years, a new group was discovered, the cedratviruses, currently composed of only two isolates and members of a putative new family, "Pithoviridae," along with previously known pithoviruses. Here we report the isolation and biological and genomic characterization of two novel cedratviruses isolated from samples collected in France and Brazil. Both viruses were isolated using *Acanthamoeba castellanii* as a host cell and exhibit ovoid particles with corks at either extremity of the particle. Curiously, the Brazilian cedratvirus is ~20% smaller and presents a shorter genome of 460,038 bp, coding for fewer proteins than other cedratviruses. In addition, it has a completely asyntenic genome and presents a lower amino acid identity of orthologous genes (~73%). Pangenome analysis comprising the four cedratviruses revealed an increase in the pangenome concomitant with a decrease in the core genome with the addition of the two novel viruses. Finally, phylogenetic analyses clustered the Brazilian virus in a separate branch within the group of cedratviruses, while the French isolate is closer to the previously reported *Cedratvirus lausannensis*. Taking all together, we propose the existence of a second lineage of this emerging viral genus and provide new insights into the biodiversity and ubiquity of these giant viruses.

IMPORTANCE Various giant viruses have been described in recent years, revealing a unique part of the virosphere. A new group among the giant viruses has recently been described, the cedratviruses, which is currently composed of only two isolates. In this paper, we describe two novel cedratviruses isolated from French and Brazilian samples. Biological and genomic analyses showed viruses with different particle sizes, genome lengths, and architecture, revealing the existence of a second lineage of this new group of giant viruses. Our results provide new insights into the biodiversity of cedratviruses and highlight the importance of ongoing efforts to prospect for and characterize new giant viruses.

KEYWORDS *Cedratvirus*, giant virus, NCLDV, new lineage, virion volume, genome length, pangenome

Viruses are the most abundant microorganisms in the biosphere and present the greatest genetic and morphological diversity among the known biological organisms (1, 2). Different groups of viruses have specific characteristics that define them as a group, such as capsid structure (e.g., icosahedral and helical) and type of genome (e.g., double-stranded DNA [dsDNA] and single-stranded RNA negative sense [ssRNA−]), which implicate differences in the life cycles and evolution of these viruses. Moreover, remarkable differences can be seen in the virion volumes and genome

Received 6 March 2018 **Accepted** 13 April 2018

Accepted manuscript posted online 25 April 2018

Citation Rodrigues RAL, Andreani J, Andrade ACSP, Machado TB, Abdi S, Levasseur A, Abrahão JS, La Scola B. 2018. Morphologic and genomic analyses of new isolates reveal a second lineage of cedratviruses. J Virol 92: e00372-18. https://doi.org/10.1128/JVI.00372-18.

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

Address correspondence to Bernard La Scola, bernard.la-scola@univ-amu.fr.

R.A.L.R. and J.A. contributed equally to this work.

lengths of viruses, exhibiting a difference of 4 orders of magnitude in the former and ranging from 1.2 kbp to 2,500 kbp in the latter (3). Despite these differences, the virion sizes and genome lengths of viruses display an allometric relationship independent of the type of capsid or genetic material, the ebolaviruses (ssRNA−) being the only exception to this scaling law described to date (3). This relationship has also been observed for some giant viruses such as mimivirus and pandoravirus (both dsDNA).

The giant viruses were discovered at the beginning of the last decade with the description of mimiviruses, revealing a new world within the virosphere (4, 5). These viruses replicate in free-living amoebas of the *Acanthamoeba* genus, although other protists have been associated with giant viruses, such as *Cafeteria roenbergensis* (6) and *Bodo saltans* (7). The discovery of viruses with gigantic particles (>500 nm) and the presence of genes considered hallmarks of the cellular world (e.g., those encoding aminoacyl-tRNA synthetases and translation factors) broke many paradigms of classical virology and reopened a hot debate about the origin and evolution of viruses (8–13). In subsequent years, other giant viruses have been isolated and characterized (14, 15), thus expanding the group of nucleocytoplasmic large DNA viruses (NCLDVs)—the proposed "Megavirales" order (16).

Among these new viruses, the pithoviruses drew attention due to their huge size (~1.5 μm) and relatively "small" genomes (~610 kbp) (17, 18), which suggest that those viruses do not fit the scaling law observed for other viruses, although no analysis has been performed in this regard to date. The same case would appear to apply to the cedratviruses, a new group of recently described viruses, with only two members characterized thus far, *Cedratvirus A11* (19), and *Cedratvirus lausannensis* (20). These viruses have an ovoid particle of about 1.0 μm and possess a circular dsDNA genome of ~590 kbp. These viruses are related to the pithoviruses but have two corks in the viral particle, one at either extremity, rather than the single one displayed by pithoviruses. Recently, it has been reported that some pithovirus particles can have complex alternative shapes and sometimes have two corks, as observed for cedratviruses (21). These viruses replicate in *Acanthamoeba* sp., entering the cells by phagocytosis. The genome is released through the cork, and an eclipse phase is established, followed by the formation of an electron-lucent viral factory, where a complex morphogenesis occurs (19, 20, 22). It is possible that there is a nuclear phase during the replication of cedratviruses, since the host nucleus remains intact during the viral cycle, although further investigation into this aspect needs to be performed (22). After ~12 h of infection, mature viral particles are released by cell lysis. The real extent of the diversity of cedratviruses is currently unknown, and the discovery of new members of this group could bring valuable information about it.

Here we describe the isolation and biological and genomic analyses of two new cedratviruses, one from samples collected in France and a second from samples collected in Brazil, which have morphological and genomic features distinct from those of the previously known cedratviruses, suggesting the existence of a second lineage among this new group of viruses. The discovery of new cedratviruses in different parts of the world reflects their ubiquity and high diversity and improves our knowledge about these viruses, thus reinforcing the importance of continuing to prospect for and biologically/genomically characterize the giant viruses.

RESULTS

New cedratviruses with different virion sizes and genome lengths. In the search for a better understanding of the diversity of giant viruses in different parts of the world, we isolated two new cedratviruses, named *Cedratvirus Zaza IHUMI* and *Brazilian Cedratvirus IHUMI*. Transmission electron microscopy (TEM) analyses revealed viruses with ovoid particles and with corks at either extremity of the particle (Fig. 1A to C; see Fig. S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses), a singular feature of cedratviruses (19, 20). Unlike pithoviruses, cedratviruses usually have two corks, although some alternative shapes with a single cork can

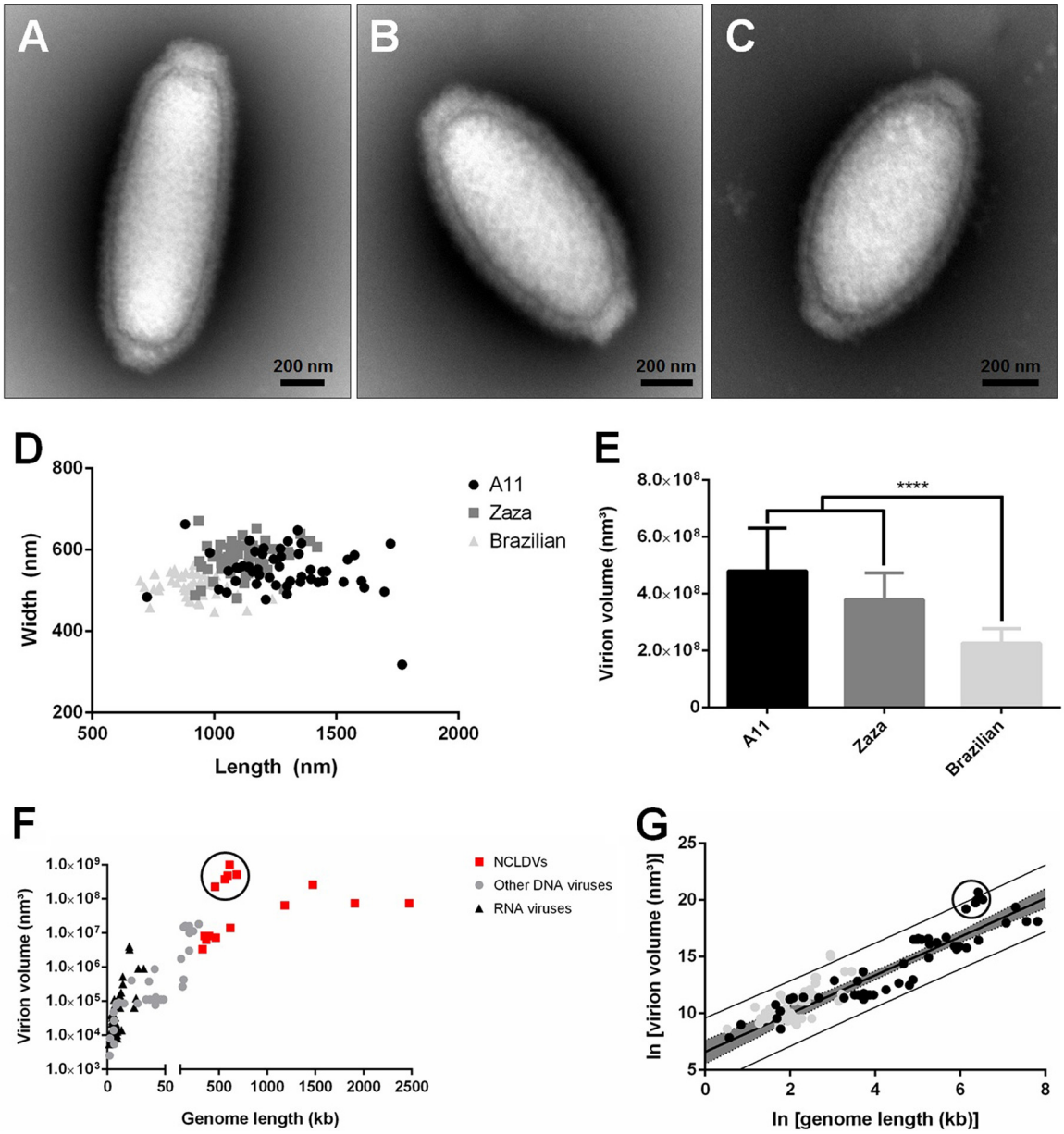


FIG 1 Morphology and volume analysis of new cedratviruses. (A to C) Negative-staining images exhibiting the characteristic ovoid shape and the presence of two corks in the particles of cedratvirus A11, cedratvirus Zaza IHUMI, and Brazilian cedratvirus IHUMI, respectively. Scale bars are indicated on each panel. (D) Length and width of 50 particles of each cedratvirus. Each point represents a single particle analyzed by using ImageJ software. (E) Volumes of different cedratviruses based on the analyses of 50 individual particles, indicating that Brazilian cedratvirus IHUMI has a significantly smaller volume than the other viruses. (F) Relationship between genome length and virion volume for different DNA and RNA viruses. Black circles highlight the pithoviruses and cedratviruses. (G) Relationship between genome length and virion volume for different viruses. The solid black line marks the linear regression between ln-ln-transformed data. The gray area represents the 95% confidence interval for the linear regression line. Black circles highlight the pithoviruses and cedratviruses. The outer gray lines represent the 95% prediction interval, within which we expect 95% of virion volume to lie for a given genome size. ****, $P < 0.0001$ (ANOVA).

be discerned, such as in pithoviruses. Moreover, as recently described for pithoviruses (21), we also observed membranous structures in empty particles in some negative-staining images (see Fig. S2 posted at the URL mentioned above). Cedratvirus Zaza IHUMI particles have a mean size of 1,110 nm (range, 921 to 1,420 nm) in length and 580 nm (range, 481 to 671 nm) in width, values closer to those observed for cedratvirus A11 (1,280 nm by 550 nm), while the Brazilian cedratvirus IHUMI particle is smaller, with the particle displaying a mean size of 910 nm (696 to 1,237 nm) in length and 510 nm

TABLE 1 Main genomic characteristics of known cedratviruses[a]

Virus	Mean particle length × width (nm)	Genome size (bp)	GC content (%)	No. of predicted proteins	Coding density (%)
Cedratvirus A11	1,280 × 550	589,068	42.6	574	78.5
Cedratvirus lausannensis	~1,000 × 500	575,161	42.8	643	83
Cedratvirus Zaza IHUMI	1,110 × 580	560,887	42.7	636	84.3
Brazilian cedratvirus IHUMI	910 × 510	460,038	42.9	533	84.3

[a]All these viruses showed ovoid, double-cork morphology, and none had tRNA.

(448 to 563 nm) in width (Fig. 1D). The difference in the particle size reflects the difference in virion volume, in that the Brazilian Cedratvirus IHUMI has the smallest volume (2.26×10^8 nm^3) among the cedratviruses analyzed, significantly smaller than cedratvirus A11 (4.8×10^8 nm^3) and cedratvirus Zaza IHUMI (3.79×10^8 nm^3) ($P < 0.0001$) (Fig. 1E). Despite these physical differences, the replication cycle of the Brazilian isolate is similar to those previously observed for other cedratviruses, exhibiting the same infection profile (see Fig. S1 posted at the URL mentioned above). The virus enters the host through phagocytosis and releases the capsid content into the cytoplasm, establishing an eclipse phase 2 h after infection. A viral factory is formed in the cytoplasm, where morphogenesis occurs, and mature virions are released 12 h after infection.

In addition to the size of the particles, the new cedratviruses have different genome lengths. Cedratvirus Zaza IHUMI has a genome of 560,887 bp coding for 636 proteins, while the Brazilian cedratvirus IHUMI has a genome of 460,038 bp coding for 533 proteins. Despite the remarkable difference in the length and numbers of predicted open reading frames (ORFs) in the genomes of the new viruses, both exhibit a circular dsDNA genome with the same coding density (84.3%) and very similar G+C contents, 42.7% and 42.9% for cedratvirus Zaza IHUMI and Brazilian cedratvirus IHUMI, respectively (Table 1). It is noteworthy that the Brazilian isolate is the smallest cedratvirus described to date and also has the smallest genome among representatives of this new group of viruses.

Cedratviruses and pithoviruses: exceptions to the allometric scaling law? The fact that the new cedratviruses exhibit different genome lengths led us to analyze the relationship between the genome length and volume size of different NCLDVs, in order to check whether they lie within the prediction interval and are thus in line with the allometric scaling law, as observed for other groups of viruses (3). We calculated the volume of 15 different viruses, including mimiviruses, marseilleviruses, pithoviruses, cedratviruses, faustovirus, kaumoebavirus, pacmanvirus, phycodnavirus, and iridovirus. The volume was calculated by considering the dimensions of viral particles resulting from previous studies using cryo-EM or negative-staining methods (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses), with the exception of the cedratviruses, for which the volume used in the analysis was the mean volume obtained from the analyses of 50 viral particles using the negative-staining method. Data concerning all other viruses were obtained from previous studies (3).

The volume of the viruses varied by 5 orders of magnitude, with porcine circovirus 1 displaying the smallest volume of the viruses under consideration (2.5×10^3 nm^3), and pithovirus sibericum presenting the largest volume (9.9×10^8 nm^3). Regarding genome length, this ranged from 1.76 kbp (porcine circovirus 1) to 2,474 kbp (pandoravirus salinus) (Fig. 1F). Considering only the volumes of the NCLDVs calculated in this study, volumes ranged from 3.31×10^6 nm^3 (chilo iridescent virus) to 9.9×10^8 nm^3 (pithovirus sibericum) (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses).

Plotting the new data on NCLDVs alongside other viruses on a log-log scale, the linear relationship is maintained ($P < 0.0001$, $R^2 = 0.83$, slope = 1.58), with values even more stringent than those previously reported (3). The vast majority of viruses fall within the 95% prediction interval, indicating that almost all viruses follow the allo-

metric scaling law for volume size and genome length, i.e., the larger the volume size of a viral particle, the longer the genome enclosure by the virus (Fig. 1G). Curiously, cedratviruses and pithoviruses are at the limit of the prediction level, with pithovirus sibericum actually outside the interval. The same was observed when considering only dsDNA viruses (data not shown). This suggests that the putative "Pithoviridae" family could be the first dsDNA group of viruses that does not conform to the allometric scaling law, along with ebolaviruses (ssRNA−). It is notable that although cedratviruses and pithoviruses appear to be exceptions to this scaling law, this appears to be true only when comparing group of viruses, since a virus with a larger volume (e.g., cedratvirus A11) has a longer genome than does a virus displaying a smaller volume, as verified for the Brazilian cedratvirus IHUMI.

Genome comparison of new cedratviruses. The cedratvirus Zaza IHUMI genome exhibit 636 genes, of which 313 (49.2%) code for proteins with no known function (hypothetical proteins). Of these, three had no hits in all searched databases and were considered ORFans (proteins that were longer than 100 amino acids and with no hits in any database). Regarding Brazilian cedratvirus IHUMI, 269 of its 533 predicted genes (50.5%) have no known function and 11 are considered to be ORFans. Among the ORFs with known functions, the presence of genes related to the metabolism of nucleic acids (e.g., those coding for DNA polymerase, DNA-dependent RNA polymerase, helicases, nucleases, DNA repair proteins) and transcription process (e.g., TFIIB initiation factor, TFIIS elongation factor, viral transcription late factor 3) was observed. Moreover, we identified 76 ankyrin repeat-containing-domain proteins in the cedratvirus Zaza IHUMI genome, while only 42 were observed in the Brazilian cedratvirus IHUMI genome. No tRNA or aminoacyl-tRNA synthetases were detected in the genomes of the new viruses. Regarding the nucleocytoplasmic virus orthologous group (NCVOG) core genes, we found some conserved genes also present in some other NCLDVs, e.g., those encoding a divergent major capsid protein (NCVOG0022), D5 helicase-primase (NCVOG0023), DNA topoisomerase II (NCVOG0037), ribonucleotide reductase (NCVOG0276 and NCVOG1353), and an mRNA capping enzyme (NCVOG1117) similar to that observed for other cedratviruses (19, 20).

Although the gene content does not exhibit significant differences at first glance, the genome organization of the Brazilian isolate is completely different from that observed for other cedratviruses, being totally asyntenic (Fig. 2). The synteny analysis revealed the presence of conserved and aligned blocks between cedratvirus Zaza IHUMI, cedratvirus A11, and cedratvirus lausannensis, while the same blocks are organized in a different orientation in the Brazilian cedratvirus IHUMI genome. Compared to the genome of the other viruses, the genome of the Brazilian isolate exhibits many inversions and rearrangements of blocks throughout its entire length. Such differences in the genomic architecture among similar viruses are observed among different lineages of mimiviruses (23) and marseilleviruses (24), which led us to consider the existence of a second lineage of cedratviruses, with Brazilian cedratvirus IHUMI being its first member.

In addition, the Brazilian cedratvirus IHUMI amino acid sequences showed lower identity than other cedratviruses (Fig. 3). The orthologous genes of the Brazilian isolate have a mean identity of 73.48% compared to cedratvirus A11, 73.6% compared to cedratvirus lausannensis, and 73.56% compared to cedratvirus Zaza IHUMI (Fig. 3A to C). In contrast, when we compared the orthologous genes from other cedratviruses to one another, we observed an amino acid identity higher than 90%, reaching 95.76% between cedratvirus lausannensis and the new isolate, cedratvirus Zaza IHUMI (Fig. 3D to F). Therefore, not only is the genomic architecture between the Brazilian isolate and the other viruses different, but also amino acid identity is considerably different, reinforcing the existence of a new lineage among the group of cedratviruses.

Pangenome and phylogenetic analyses of cedratviruses. The pangenome analysis of the cedratviruses isolated thus far revealed an increase in the pangenome content with the addition of a gene repertoire by way of the new viruses described in

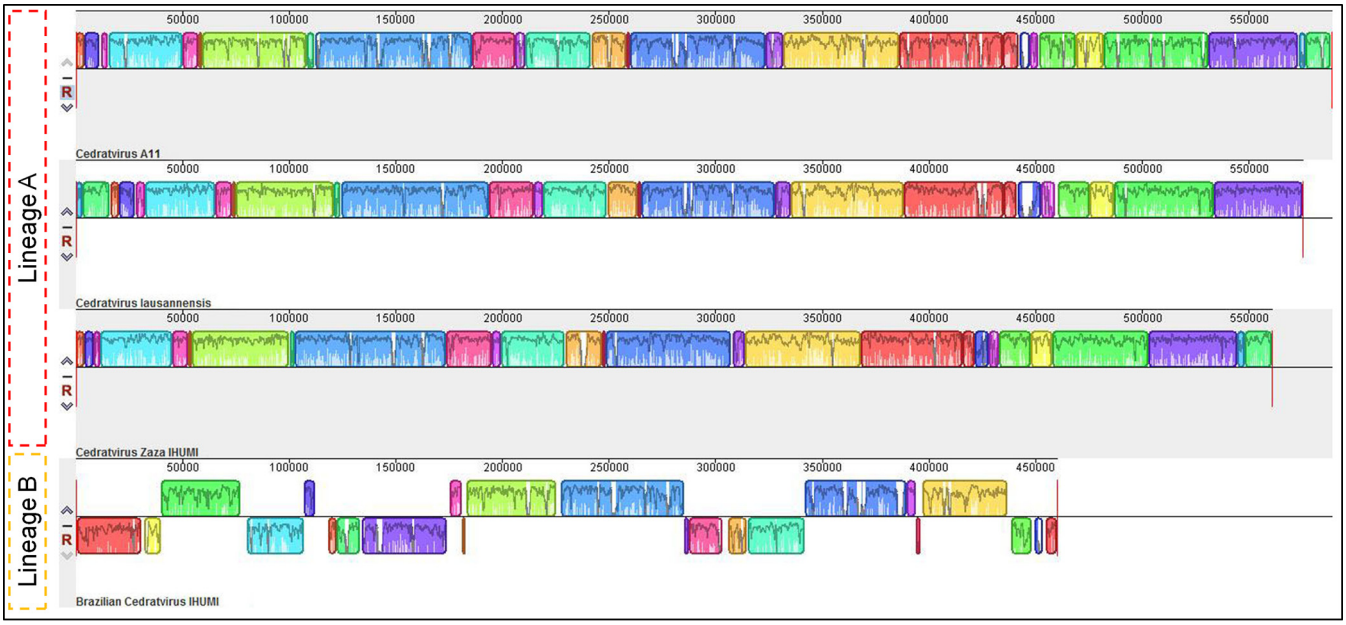


FIG 2 Genome synteny analysis. Schematic genome alignment diagram obtained using the Mauve software package. The analysis was performed using the genome of cedratvirus A11 (NC_032108.1) and cedratvirus lausannensis (LT907979.1), besides the genome sequences of the new isolates. The blocks illustrated above the *x* axis are in the positive strand (forward sense), while blocks below the *x* axis are in the negative strand (reverse sense).

this study. A total of 2,386 proteins were grouped into 821 clusters of orthologous genes (COGs) (Fig. 4), including 613 COGs comprising genes for at least two proteins of different cedratviruses. The size of the pangenome content displayed a continuous increase with the addition of the two new viruses, including an increase of 61 new COGs with the addition of the Brazilian cedratvirus IHUMI, even though this virus presented a genome coding for fewer proteins than the other viruses. Furthermore, it is the virus that presents the greatest number of unique COGs (numbering 72), while the others present only 59 (cedratvirus lausannensis), 47 (cedratvirus Zaza IHUMI), and 30 (cedratvirus A11) unique COGs (see Fig. S3 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses). However, the most remarkable point is the existence of a break in the curve of the core genome content when genes of the Brazilian isolate are added (−102), reaching a total of 386 COGs for this proposed viral genus (Fig. 4). These data suggest that different lineages of cedratviruses could contribute to a slight increase in the pangenome and could share a reduced core gene set.

To better understand the evolutionary relationship between the new cedratviruses and other members of the proposed Megavirales order, we performed phylogenetic analyses based on different NCLDV genes (NCVOGs) including those coding for the family B DNA polymerase (NCVOG0038) (Fig. 5), the major capsid protein (NCVOG0022), the DNA-dependent RNA polymerase subunit 1 (NCVOG0274), and the VV-A18 helicase (NCVOG0508) (see Fig. S4 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses). Moreover, we performed additional phylogenetic analyses using the D6/D11 helicase (NCVOG0031), DNA repair exonuclease (NCVOG0308), Flap endonuclease (NCVOG1060), and ATP-dependent DNA ligase (NCVOG0034), focusing on the cedratviruses and closer viral groups, i.e., marseilleviruses and irido/ascoviruses (see Fig. S5 posted at the URL mentioned above). Phylogenetic trees recurrently clustered the new isolates alongside previously described cedratviruses, pithoviruses, and orpheovirus. Furthermore, all trees based on the core genes showed the cedratvirus Zaza IHUMI as being closer to cedratvirus lausannensis and cedratvirus A11 and the Brazilian isolate being in a branch distant

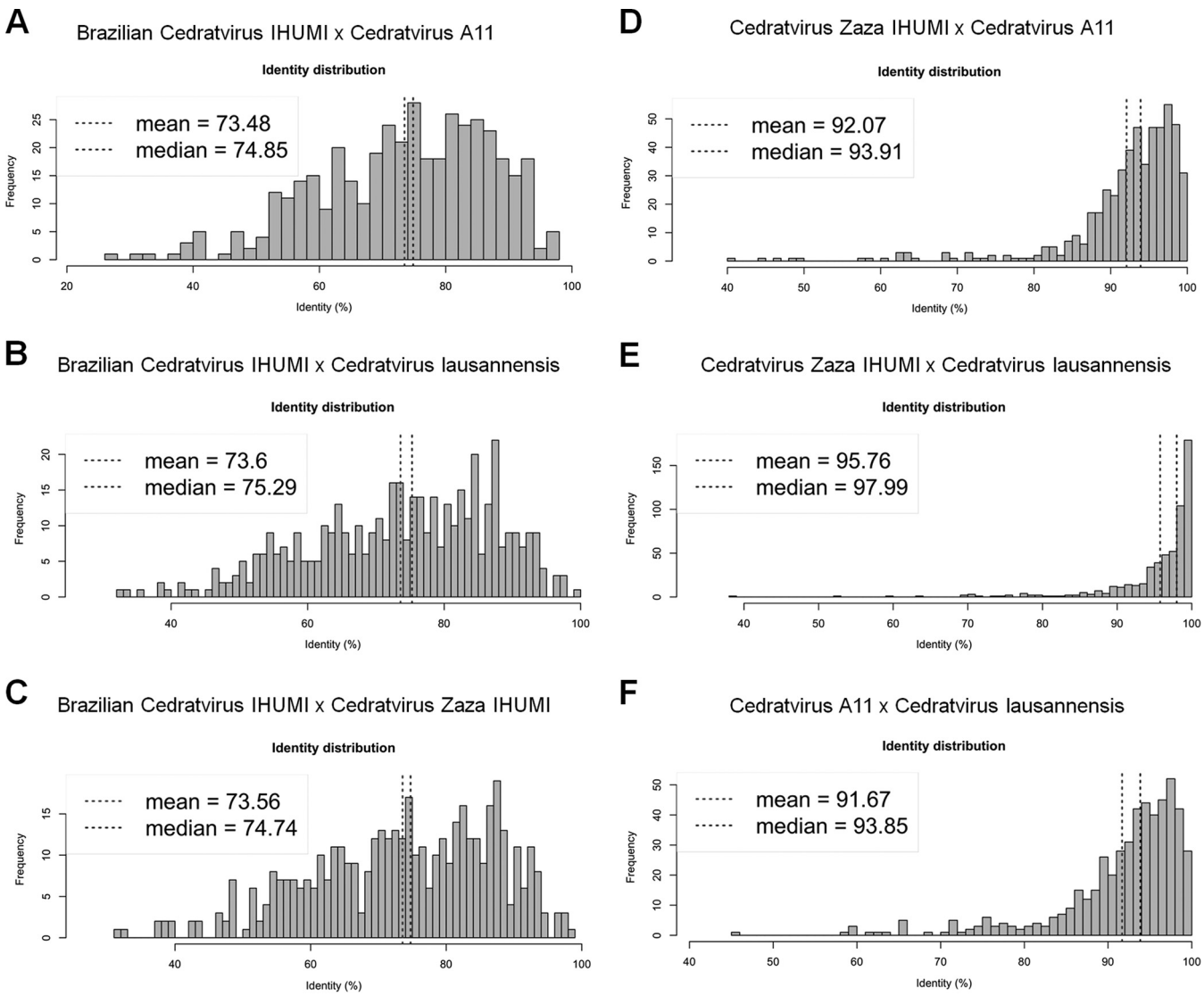


FIG 3 Average amino acid identity. In this analysis, estimates were reached using both best hits (one-way AAI) and reciprocal best hits (two-way AAI) between two data sets of proteins from different cedratviruses. Plots A to C demonstrate the amino acid comparisons between Brazilian cedratvirus IHUMI and other cedratviruses; plots D and E compare cedratvirus Zaza IHUMI and previously known cedratviruses; plot F compares cedratvirus A11 and cedratvirus lausannensis.

from the other cedratviruses with a bootstrap value of >90, corroborating the existence of a new lineage among cedratviruses. Finally, the putative "Pithoviridae" family is clustered along with marseilleviruses or irido/ascoviruses depending of the gene used, the tree topology not being always congruent. An in-depth phylogenetic analysis must be performed to better establish the phylogenetic relationship among these groups of giant viruses.

DISCUSSION

The isolation of new giant viruses associated with biological and genomic analyses has significantly contributed to broadening our understanding of the diversity, ecology, and evolution of this complex group of viruses. The discovery of pithoviruses (17, 18) and cedratviruses (19, 20) drew particular attention, since these viruses exhibit very large particles constraining relatively short genomes, forming a putative novel viral family among the group of NCLDVs. In this study, we describe the isolation and the biological and genomic analyses of two new members of this group, providing new insights into the biodiversity and evolution of these viruses.

The analyses performed in this study revealed two new viruses with significant

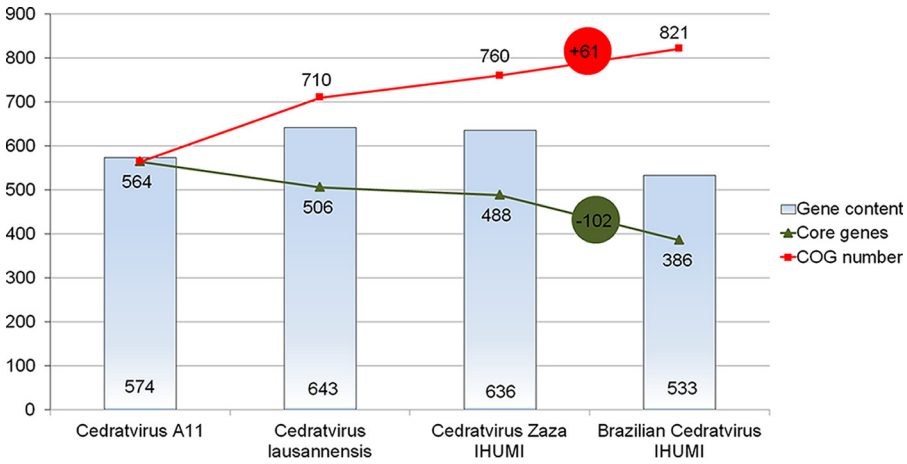


FIG 4 Pangenome (red line) and core genome (green line) sizes of cedratviruses. Numbers at the base of the column refer to the number of genes carried by each virus strain. Numbers at line nodes represent the cumulative COG numbers after the inclusion of a new genome. Numbers in (red and green) circles demonstrate the variation of COGs after the inclusion of the Brazilian cedratvirus IHUMI (proposed new lineage).

structural differences, both physical and genomic. Although the cedratvirus Zaza IHUMI exhibits a particle size and genome length similar to those of other cedratviruses that have been described, analysis of the Brazilian cedratvirus IHUMI revealed a virus with particles that were smaller (~20%) than those of the other viruses of the same group and a considerably smaller genome. By analyzing the relationship between the virion volume and the genome length of viruses, including those from different groups of giant viruses, we noticed that the majority of viruses fall into the allometric scaling law and, curiously, the pithoviruses and cedratviruses are at the limits of the prediction interval. This suggests that these viruses might be exceptions to this scaling law. Since we considered only data from comparable imaging methods (i.e., cryo-EM and negative staining [25, 26]) to calculate the volumes of giant viruses, only a few viruses were analyzed. It is possible that with new, forthcoming structural data on viruses, particularly on giant viruses, it may be discovered that the pithoviruses and cedratviruses definitively fall outside the prediction interval. Indeed, when the virion size data for other amoebal giant viruses (e.g., mimiviruses and marseilleviruses) from TEM images were considered in our analysis, the members of the putative "Pithoviridae" did not fit with the allometric scaling law (data not shown). It is notable that, along with ebolaviruses (*Filoviridae* family), the members of the putative "Pithoviridae" family are the only known viruses that display a massive particle but a "small" genome. Such features raise important questions about what is inside these viral particles. A recent study comparing the internal density of pithoviruses' and mimiviruses' particles demonstrated that the former viruses have three-quarters of the internal density of the latter, suggesting that the pithoviruses may carry macromolecules other than nucleic acids inside the particles (21). The same would appear to be the case for the cedratviruses, but further studies are needed to define exactly which macromolecules could be carried by those viruses.

The fact that the Brazilian isolate has a smaller genome is equally curious. Similar to other cedratviruses, this new virus exhibits only a few repeat zones throughout the genome (data not shown), and we identified the presence of genes also present in other cedratviruses, such as those coding for polymerases, helicases, nucleases, etc. Among the genes with known function, we noticed differences mainly in the quantity of those coding for proteins containing repeat domains, especially coding for ankyrin repeat motifs, as Brazilian cedratvirus IHMU (a virus with a smaller genome) has fewer genes of these category than do other cedratviruses. This is in

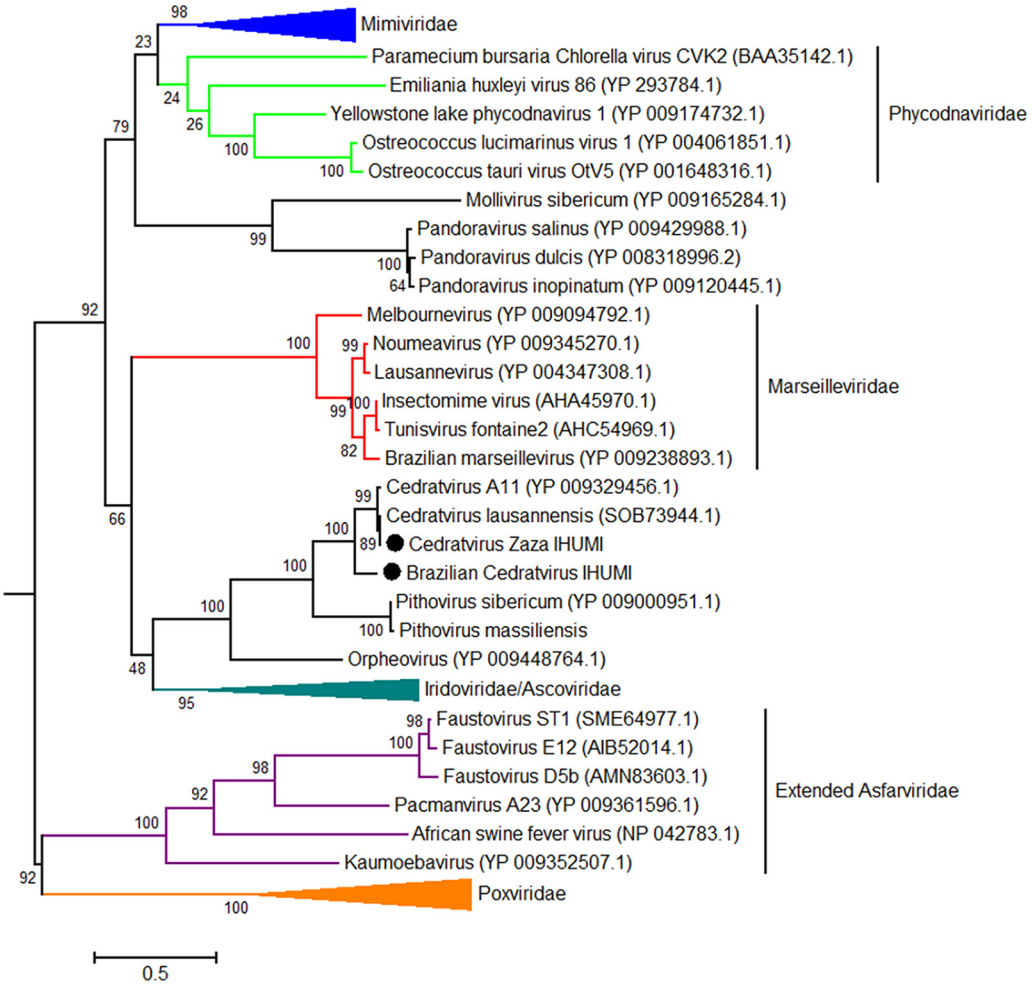


FIG 5 Phylogenetic tree based on DNA polymerase B amino acid sequences of nucleocytoplasmic large DNA viruses (NCLDVs). The tree was constructed using MEGA version 6.0, applying the maximum likelihood method and the JTT model of evolution with 1,000 bootstrap replicates. Colors indicate viral families: blue was used for *Mimiviridae*; green for *Phycodnaviridae*, red for *Marseilleviridae*, navy blue for *Iridoviridae/Ascoviridae*, purple for extended *Asfarviridae*, and orange for *Poxviridae*. The new cedratviruses are highlighted with black circles. The scale bar indicates the rate of evolution.

accordance with the recent analysis conducted by Shukla and colleagues, wherein they demonstrated that in different groups of giant viruses infecting amoebae, the quantity of this class of genes is proportional to the length of the genome (27), which has also been observed for some intracellular bacteria (28). These analyses also seem to apply to viruses within the same group, such as the cedratviruses described here. Taking this into account, it is possible that the Brazilian cedratvirus IHUMI underwent different selective pressures, thus contributing to a different evolutionary history. This would be in accordance with our proposed new lineage within the cedratviruses. Such a proposal is supported by the observation of a completely different genomic architecture between the Brazilian cedratvirus IHUMI and the other viruses, in addition to a significant difference in the amino acid identity of orthologous genes, similar to that observed for members of the *Mimiviridae* and *Marseilleviridae* families (23, 24). In addition, this virus has more exclusive COGs and contributes to an increase in the pangenome with 61 new COGs and, even more strikingly, with the reduction of the core genome by 102 COGs. Finally, phylogenetic analyses based on different core genes of giant viruses clearly clustered the Brazilian isolate in a separate branch from other cedratviruses, therefore reinforcing the existence of a lineage "B" among this new group of viruses.

It is still premature to dive deep into the evolutionary history of cedratviruses, but it is possible that they underwent an accordion-like model of evolution as observed for other giant viruses (29), although new analysis must be performed to confirm this hypothesis. In any case, it is clear that this new, expanding group of viruses deserves attention, and new structural and evolutionary analyses could help to solve some unanswered questions around them.

MATERIALS AND METHODS

Virus isolation, production, and purification. Two novel cedratviruses were isolated by the coculture method as previously described (30), one from an *Alpova* sp. (*Basidiomycota*, *Paxillaceae* family) homogenate in sterile distilled water collected near Toulon, France, and the other from a water sample supplemented with bio-floc, collected in Belo Horizonte, Brazil. The former virus was isolated at the Institut-Hospitalo-Universitaire (IHU) Méditerrannée Infection at Marseille, France, and was named *Cedratvirus Zaza IHUMI*, while the second was isolated in the Laboratürio de Vírus of UFMG at Belo Horizonte, Brazil. The Brazilian isolate was then sent to IHU for further production, genome sequencing, and analysis and was given the name of *Brazilian Cedratvirus IHUMI*. For multiplication of the viruses, *Acanthamoeba castellanii* (strain Neff [ATCC 30010]) was cultured in a 150-cm^2 cell culture flask with 50 ml of a peptone-yeast extract-glucose (PYG) medium at 28°C. When the flasks contained a fresh monolayer of *A. castellanii*, the PYG medium was replaced by starvation medium (TS). The amoebas were then infected with the isolated virus, and the flasks were kept at 30°C for 72 h. The cell lysates were then collected and centrifuged at 400 $\times$ *g* for 10 min to remove amoeba debris. The supernatants were then centrifuged at 6,500 $\times$ *g* for 15 min at 4°C, and the pellets were suspended in Page's amoeba saline (PAS) solution. This process was repeated twice. The pellets were suspended in 3 ml of phosphate-buffered saline (PBS) solution, added to a sucrose cushion (50%), and centrifuged at 16,000 $\times$ *g* for 15 min at 4°C. The final pellets were suspended in PAS solution.

Characterization of the replicative cycle. In order to study possible differences in the replicative cycle of Brazilian cedratvirus IHUMI, ultrathin sections of infected amoebas were evaluated under TEM and a comparative one-step growth curve was performed. For the microscopy analysis, *A. castellanii* cells were infected with Brazilian cedratvirus IHUMI at a multiplicity of infection (MOI) of 0.01 for 24 h in PYG medium (asynchronous cycle). The cells were then collected and fixed 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 under transmission electron microscopy (Spirit Biotwin FEI, 120 kV). For the one-step growth curves, *A. castellanii* cells were infected with different cedratviruses at an MOI of 10 in TS medium in 24-well microplates. After 30 min of adsorption, the inoculum was removed, and fresh medium was added. The cell and supernatant were collected at different time points and further titrated using the endpoint method (31). The experiment was performed in duplicate.

Virus particle morphometry and volume calculation. For particle morphometry, negative staining was performed on the fixed supernatant from coculture. A total of 5 μl was deposited onto the glow-discharged grid for 20 min at room temperature. The dried grid was contrasted with a small drop of 1% ammonium molybdate for 5 s, and the grid was then observed using a Tecnai G20 electron microscope (FEI, Germany) operated at 200 kV. At least 50 particles of each virus were analyzed using ImageJ software (32). For the volume calculation of cedratvirus particles, we employed the formula for spheroid particles as previously described for ovoid viruses (3), $V = 4/3 \times \pi a^2 c$, where V is the volume of the viral particle, a is the equatorial radius of the spheroid, c is the distance from the center to the pole along the symmetry axis, and π is a constant. The data used for the volume calculation of other NCLDVs were obtained from previous publications, considering data only from cryo-electron microscopy or transmission electron microscopy negative-staining data (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses). For icosahedral viruses, we used the formula for spherical particles, $V = 4/3 \times \pi r^3$, employing a strategy described elsewhere (3), where r is one-half of the diameter of the virus capsid. For other viruses, we kept the volume data previously calculated by Cui and colleagues (3).

Statistical analysis. We used analysis of variance (ANOVA) to compare the virion volumes of different cedratviruses and linear regression between the natural logarithm of genome length and the natural logarithm of virion volume to test whether the allometric relationship previously described for other viruses (3) also applied to giant viruses of amoebas, which had not previously been evaluated. The statistical analyses were performed by using GraphPad Prism 6.0.

DNA extraction and genome sequencing and assembly. The genomes of the new cedratviruses were extracted using the automated EZ1 virus minikit v.2 (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. DNA was quantified using a Qubit assay with the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 30.3 ng/μl (cedratvirus Zaza IHUMI) and 16 ng/μl (Brazilian cedratvirus IHUMI). DNA was sequenced using MiSeq Technology (Illumina Inc., San Diego, CA, USA) with the paired-end application. DNA was barcoded in order to be mixed with other projects for the Nextera XT DNA sample prep kit (Illumina).

To prepare the paired-end library, dilution was performed to yield 1.0 ng of each genome as input. The "tagmentation" step fragmented and tagged the DNA. Limited-cycle PCR amplification (12 cycles) then completed the tag adapters and introduced dual-index barcodes. The library profile was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc., Santa Clara, CA, USA) with a DNA high-

sensitivity LabChip, and the fragment size was estimated to 1.5 kb. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, USA), the libraries were then normalized on specific beads according to the Nextera XT protocol (Illumina). Normalized libraries were pooled for sequencing on the MiSeq. Automated cluster generation and paired-end sequencing with dual index reads were performed in a single 39-hour run in 2× 250 bp. A total of 2.8 Gb of information was obtained from a 277,000/mm² cluster density in the first run with a cluster passing quality control filters of 98.2% (5,333,000 passed filtered clusters). Within this run, the index representation for cedratvirus Zaza IHUMI was determined to be 2.18%. The 149,880 paired-end reads were trimmed and filtered according to the read qualities. Additionally, a total of 7.5 Gb of information was obtained from an 802,000/mm² cluster density in the second run with a cluster passing quality control filters of 96.4% (14,444,000 clusters). Within this run, the index representation for Brazilian cedratvirus IHUMI was determined to be 8.75%. The 1,264,356 paired-end reads were filtered according to the read qualities.

In addition, a run was performed through the MinIon Oxford Nanopore for the Brazilian isolate. The Oxford Nanopore approach was performed on 1D genomic DNA sequencing for the MinIon device using the SQK-LSK108 kit. A library was constructed from 1.5 μg genomic DNA without fragmentation and end repair. Adapters were ligated to both ends of the genomic DNA. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, USA), the library was quantified using a Qubit assay with the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA). An amount of 136.8 ng, adapted and tethered as a library, was loaded on the flow cell via the SpotON port. A total of 1,110 active pores were detected for sequencing, and the WIMP workflow was chosen for live bioinformatic analysis. After 4 h and 40 min of run time, the EPI2ME software led to 6,299 classified reads of the Brazilian cedratvirus IHUMI of the 98,601 analyzed reads, with an average length of 2.6 kb.

The sequence reads were assembled *de novo* using the CLC Genomics Workbench v7.5 (http://www.clcbio.com/blog/clc-genomics-workbench-7-5/) for the cedratvirus Zaza IHUMI and hybridSPAdes (33) for the Brazilian cedratvirus IHUMI.

Study of genome organization and genome annotation. Open reading frames were predicted by GeneMarkS (34), and the draft genomes were deposited in NCBI. The tRNA genes were searched using the tRNAscan-SE and ARAGORN software (35, 36). Predicted proteins of fewer than 50 amino acids in length were discarded. A BLASTp search against the NCBI nonredundant (nr) protein sequence database was performed on 5 January 2018. Homology was considered significant if the E value was lower than 1×10^{-3}. A BLASTp search was also computed with the same parameters against the clusters of orthologous groups (COGs) of proteins of the nucleocytoplasmic large DNA virus (known as NCVOGs) (37). In addition, we searched for conserved domains and putative functions of predicted proteins using the online InterProScan software, version 66.0 (https://www.ebi.ac.uk/interpro/search/sequence-search). In addition, predicted ORFs ranging from 50 to 99 amino acids were submitted for tridimensional folding analyses using Phyre2 (38). Proteins ranging from 50 to 99 amino acids in length were discarded if they exhibited no hits either against the BLASTp or against the InterProScan searches or if they exhibited abnormal folding as modeled by Phyre2. Those proteins that were longer than 100 amino acids and with no hits in any database were kept and referred to as ORFans. Finally, the genome annotation was manually revised and curated.

Comparative genomic and pangenome analysis. The genome synteny among cedratviruses was checked using the Mauve program (39) with default parameters. The Proteinortho tool (40) was used to identify orthologous gene sequences based on the reciprocal best hit shared by cedratviruses using an amino acid sequence identity of 30% and sequence coverage of 60% as thresholds. The average amino acid identity (AAI) calculator tool (41) was used to compare identity between orthologous genes from cedratvirus isolates. To estimate the size of the pangenome, their predicted proteins were clustered using the Proteinortho tool (40), applying the same criteria as those given above. We also described pangenome and core genes size variation by stepwise inclusion of each new virus annotation in the pairwise comparisons of the gene contents of the available cedratvirus genome sequences.

Phylogenetic analysis. Phylogenetic reconstructions were based on individual alignment of distinct genes, namely, those encoding the DNA polymerase B family, the DNA-dependent RNA polymerase subunit 1, the VV-A18 helicase, the major capsid protein, the D6/D11 helicase, the Flap endonuclease, the ATP-dependent DNA ligase, and the DNA repair exonuclease. Amino acid sequences were aligned using the Muscle software (42). Phylogenetic trees were built using the MEGA6 software (43), the Jones-Taylor-Thornton (JTT) model for amino acid substitution, and the maximum likelihood method with 1,000 bootstrap replicates.

Accession number(s). Sequences for the draft genomes were deposited in NCBI under the accession numbers LT994651 (Brazilian cedratvirus IHUMI) and LT994652 (cedratvirus Zaza IHUMI).

ACKNOWLEDGMENTS

We thank our colleagues at the Laboratório de Vírus and Center of Microscopy of UFMG for their excellent technical support. Thanks also go to Enora Tomei and Caroline Michele for genome sequencing and to Emeline Baptiste for genome submission.

This work was funded by CNPq, FAPEMIG, CAPES-COFECUB. J.S.A. is a CNPq researcher. 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, French National Research Agency) (reference: Méditerranée Infection 10-IAHU-03).

R.A.L.R. and J.A. performed genomic analyses and wrote the paper; R.A.L.R. performed morphological and phylogenetic analyses; J.A., S.A., A.C.S.P.A., and T.B.M. performed virus isolation and initial biological characterization; A.L., J.S.A., and B.L.S. conceived the study and revised the manuscript.

REFERENCES

1. Suttle CA. 2005. Viruses in the sea. Nature 437:356–361. https://doi.org/10.1038/nature04160.
2. Hulo C, De Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, Le Mercier P. 2011. ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Res 39(Database issue):D576–D582. https://doi.org/10.1093/nar/gkq901.
3. Cui J, Schlub TE, Holmes EC. 2014. An allometric relationship between the genome length and virion volume of viruses. J Virol 88:6403–6410. https://doi.org/10.1128/JVI.00362-14.
4. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie J-M, Raoult D. 2003. A giant virus in amoebae. Science 299:2033. https://doi.org/10.1126/science.1081867.
5. Colson P, La Scola B, Levasseur A, Caetano-Anollés G, Raoult D. 2017. Mimivirus: leading the way in the discovery of giant viruses of amoebae. Nat Rev Microbiol 15:243–254. https://doi.org/10.1038/nrmicro.2016.197.
6. Fischer MG, Allen MJ, Wilson WH, Suttle CA. 2010. Giant virus with a remarkable complement of genes infects marine zooplankton. Proc Natl Acad Sci U S A 107:19508–19513. https://doi.org/10.1073/pnas.1007615107.
7. Deeg CM, Chow C-ET, Suttle CA. 2018. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. Elife 7:e33014. https://doi.org/10.7554/eLife.33014.
8. Raoult D, Forterre P. 2008. Redefining viruses: lessons from Mimivirus. Nat Rev Microbiol 6:315–319. https://doi.org/10.1038/nrmicro1858.
9. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie J-M. 2004. The 1.2-megabase genome sequence of Mimivirus. Science 306:1344–1350. https://doi.org/10.1126/science.1101485.
10. Moreira D, López-García P. 2009. Ten reasons to exclude viruses from the tree of life. Nat Rev Microbiol 7:306–311. https://doi.org/10.1038/nrmicro2108.
11. Raoult D. 2009. There is no such thing as a tree of life (and of course viruses are out!). Nat Rev Microbiol 7:615. https://doi.org/10.1038/nrmicro2108-c6.
12. Koonin EV, Krupovic M, Yutin N. 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. Ann N Y Acad Sci 1341:10–24. https://doi.org/10.1111/nyas.12728.
13. Nasir A, Caetano-Anollés G. 2015. A phylogenomic data-driven exploration of viral origins and evolution. Sci Adv 1:e1500527. https://doi.org/10.1126/sciadv.1500527.
14. 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.
15. Colson P, La Scola B, Raoult D. 2017. Giant viruses of amoebae: a journey through innovative research and paradigm changes. Annu Rev Virol 4:61–85. https://doi.org/10.1146/annurev-virology-101416-041816.
16. Colson P, De Lamballerie X, Fournous G, Raoult D. 2012. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 55:321–332. https://doi.org/10.1159/000336562.
17. 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 J-M. 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.
18. Levasseur A, Andreani J, Delerce J, Bou Khalil J, Robert C, La Scola B, Raoult D. 2016. Comparison of a modern and fossil Pithovirus reveals its genetic conservation and evolution. Genome Biol Evol 8:2333–2339. https://doi.org/10.1093/gbe/evw153.
19. Andreani J, Aherfi S, Khalil JYB, 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 8:1–11. https://doi.org/10.3390/v8110300.
20. Bertelli C, Mueller L, Thomas V, Pillonel T, Jacquier N, Greub G. 2017. Cedratvirus lausannensis—digging into Pithoviridae diversity. Environ Microbiol 19:4022–4034. https://doi.org/10.1111/1462-2920.13813.
21. Okamoto K, Miyazaki N, Song C, Maia FRNC, Reddy HKN, Abergel C, Claverie JM, Hajdu J, Svenda M, Murata K. 2017. 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 7:13291. https://doi.org/10.1038/s41598-017-13390-4.
22. Silva LKDS, Andrade ACDSP, Dornas FP, Rodrigues RAL, Arantes T, Kroon EG, Bonjardim CA, Abrahaõ JS. 2018. Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. Sci Rep 8:4000. https://doi.org/10.1038/s41598-018-22398-3.
23. Assis FL, Franco-Luiz APM, dos Santos RN, Campos FS, Dornas FP, Borato PVM, Franco AC, Abrahao JS, Colson P, La Scola B. 2017. Genome characterization of the first mimiviruses of lineage C isolated in Brazil. Front Microbiol 8:2562. https://doi.org/10.3389/fmicb.2017.02562.
24. Dornas FP, Assis FL, Aherfi S, Arantes T, Abrahão JS, Colson P, La Scola B. 2016. A Brazilian marseillevirus is the founding member of a lineage in family marseilleviridae. Viruses 8(3):76. https://doi.org/10.3390/v8030076.
25. Hoenger A, Aebi U. 1996. 3-D reconstructions from ice-embedded and negatively stained biomacromolecular assemblies: a critical comparison. J Struct Biol 117:99–116. https://doi.org/10.1006/jsbi.1996.0075.
26. De Carlo S, Harris JR. 2011. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. Micron 42:117–31. https://doi.org/10.1016/j.micron.2010.06.003.
27. Shukla A, Chatterjee A, Kondabagil K. 2018. The number of genes encoding repeat domain-containing proteins positively correlates with genome size in amoebal giant viruses. Virus Evol 4:1–11. https://doi.org/10.1093/ve/vex039.
28. Moliner C, Fournier PE, Raoult D. 2010. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol Rev 34:281–94. https://doi.org/10.1111/j.1574-6976.2009.00209.x.
29. Filée J. 2015. Genomic comparison of closely related giant viruses supports an accordion-like model of evolution. Front Microbiol 6:593. https://doi.org/10.3389/fmicb.2015.00593.
30. Khalil JYB, Robert S, Reteno DG, Andreani J, Raoult D, La Scola B. 2016. High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. Front Microbiol 7:26. https://doi.org/10.3389/fmicb.2016.00026.
31. Reed LJ, Muench H. 1938. A simple method of estimating fifty per cent endpoints. Am J Hyg 27:493–497.
32. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089.
33. Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. HybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics 32:1009–1015. https://doi.org/10.1093/bioinformatics/btv688.
34. Besemer J. 2001. 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.
35. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res 33(Web Server issue):W686–W689.
36. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.
37. 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.
38. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.
39. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. https://doi.org/10.1101/gr.2289704.

40. Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12:124. https://doi.org/10.1186/1471-2105-12-124.
41. Rodriguez-R LM, Konstantinidis KT. 2014. Bypassing cultivation to identify bacterial species. Microbe Mag 9:111–118. https://doi.org/10.1128/microbe.9.111.1.
42. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113.
43. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.

SCIENTIFIC REPORTS

OPEN

Cedratvirus getuliensis replication cycle: an in-depth morphological analysis

Received: 9 January 2018
Accepted: 22 February 2018
Published online: 05 March 2018

Ludmila Karen dos Santos Silva[1], Ana Cláudia dos Santos Pereira Andrade[1], Fábio Pio Dornas[1,2], Rodrigo Araújo Lima Rodrigues[1], Thalita Arantes[1], Erna Geessien Kroon[1], Cláudio Antônio Bonjardim[1] & Jônatas Santos Abrahão[1]

The giant viruses are the largest and most complex viruses in the virosphere. In the last decade, new members have constantly been added to this group. Here, we provide an in-depth descriptive analysis of the replication cycle of Cedratvirus getuliensis, one of the largest viruses known to date. We tracked the virion entry, the early steps of virus factory and particles morphogenesis, and during this phase, we observed a complex and unique sequential organization of immature particle elements, including horseshoe and rectangular compartments, revealed by transverse and longitudinal sections, respectively, until the formation of the final ovoid-shaped striped virion. The genome and virion proteins are incorporated through a longitudinal opening in the immature virion, followed by the incorporation of the second cork and thickening of the capsid well. Moreover, many cell modifications occur during viral infection, including intense membrane trafficking important to viral morphogenesis and release, as evidenced by treatment using brefeldin A. Finally, we observed that Cedratvirus getuliensis particles are released after cellular lysis, although we obtained microscopic evidence that some particles are released by exocytosis. The present study provides new information on the unexplored steps in the life cycle of cedratviruses.

The study of giant viruses has been intensified after the isolation of Acanthamoeba polyphaga mimivirus, a virus of outstanding dimensions, capable of infecting amoebas of the genus *Acanthamoeba*[1]. Since then, the intense prospection and improvement of isolation techniques has made possible the discovery of new viruses[2,3]. The presence of these viruses has been observed in rather diverse environments, such as water, soil, sewage, and clinical samples, as well as in extreme environments, including permafrost and soda lakes, for example[4–6]. These discoveries have revealed a wide diversity and variety of species not previously observed in the virosphere, challenging the concepts and paradigms concerning the canonical definition of viruses[7]. Currently, the International Committee of Taxonomy of Viruses (ICTV) officially recognizes two families of giant virus of amoebas: *Mimiviridae* and *Marseilleviridae*. In addition to these families, other giant viruses (not assigned yet) have been isolated, such as Faustovirus and Kaumoebavirus, the first giant viruses described to replicate in *Vermoamoeba vermiformes*[8,9]. The tupanviruses, recently isolated from Brazilian environments, present a complex virion structure, with a mimivirus-like capsid attached to a long tail, and these viruses replicate in a broad range of protists (unpublished data). Other isolated viruses, such as Pandoravirus, Pithovirus, Mollivirus and Cedratvirus, also have atypical virion morphologies, exhibiting amphora-shaped, spherical or ovoid structures[4,6,10,11].

Among these viruses, the cedratvirus has an ovoid viral particle, morphologically similar to that of pithovirus but presenting two corks, one at each apex[4,10]. The first Cedratvirus, A11, was isolated from environmental samples from Algeria[10]. Then, a second isolate, Cedratvirus lausannensis, was recovered from a water treatment plant in Morsang-sur-Seine, France[12]. Through an extensive prospective study, we isolated the first cedratvirus from Brazil, named Cedratvirus getuliensis. Although studies on the prospection of giant viruses have advanced over the years, enabling the isolation of new viruses, information regarding their biology remains scarce. In the present study, we present an in-depth investigation of the replication cycle of Cedratvirus getuliensis (C. getuliensis). Through transmission electron microscopy and biological assays using different pharmacological

[1]Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil. [2]Centro de Microscopia da Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil. Correspondence and requests for materials should be addressed to J.S.A. (email: jonatas.abrahao@gmail.com)

inhibitors, we elucidated different steps of the replication cycle. We provided the first evidence of a complex and unique sequential organization of immature particles elements, including transverse-sectioned horseshoe and longitudinal-sectioned rectangular compartments, until the formation of the final striped, ovoid-shaped virion. Moreover, many cell modifications occur during viral infection, raising questions about the role of some organelles during the replication of Cedratvirus getuliensis. Amorphous particles were observed in many cells, similar to those previously observed for Pithovirus, but these particles were homogeneously diffused throughout the host cytoplasm, suggesting that deformed particles are naturally formed by Cedratvirus getuliensis. Finally, we observed that Cedratvirus getuliensis particles are released after cellular lysis, although we obtained microscopy evidence that some particles are released by exocytosis. These results provide new information on the unexplored steps in the life cycle of cedratviruses.

Material and Methods

Virus isolation, cell culture, production and titration. Cedratvirus getuliensis was previously isolated in 2017 from sewage samples collected in the city of Itaúna, Minas Gerais, Brazil. After isolation, the virus genome was sequenced, and subsequent bioinformatics analyses were developed; we observed high homology and synteny among the genomes of Cedratvirus getuliensis and other Cedratviruses (in preparation). For co-culture and isolation procedures, *Acanthamoeba castellanii* cells (ATCC 30010) were cultivated in Peptone-yeast extract with glucose (PYG)[13] medium supplemented with 25 mg/ml amphotericin B (Fungizone; Cristalia, São Paulo, Brazil), 500 U/ml penicillin (Schering-Plough, Brazil) and 50 mg/ml gentamicin (Schering-Plough, Brazil). A total of $7 \times 10E6$ cells was infected with *C. getuliensis* at a multiplicity of infection (MOI) of 0.01 and incubated at 32 °C. After the appearance of a cytopathic effect, the cells and supernatants were collected, with sterile serological pipettes, stored in conic sterile tubes and the viruses were subsequently purified through ultracentrifugation with a 40% sucrose cushion at 36,000 *g* for 1 h. After purification, the viruses were serially diluted, and multiple replicate samples of each dilution were inoculated into *A. castellanii* monolayers. After 72–96 h of incubation, the amoebas were analyzed to determine whether infection occurred. Based on these data, the virus titers were determined using the endpoint method[13,14].

Entry and traffic membrane assays. In these experiments, we first evaluated the primary mechanism used by C. getuliensis to enter *A. castellanii* cells. For that we used different chemical inhibitors in order to investigate different endocytic pathways commonly explored by viral particles to enter in host cells, such as cytochalasin D – a phagocytosis inhibitor, chloroquine – clathrin and caveolin -dependent of acidification pathways inhibitors, and 5-(N-ethyl-N-isopropyl) amiloride (EIPA) – a specific macropinocytosis inhibitor. Cytochalasin D and chloroquine had already been confirmed as inhibitors of endocytic pathways in *Acanthamoeba*. However, the micropinocytosis inhibition effect induced by EIPA (observed in other systems) remains to be molecularly investigated in *Acanthamoeba*. A total of 5×10^5 *A. castellanii* cells was pre-treated with 2 μM of cytochalasin (Sigma-Aldrich, United States), 100 μM of chloroquine (Sigma-Aldrich, United States) or 1 μM of EIPA (Sigma-Aldrich, United States). The cytotoxicity of the inhibitors was tested in *Acanthamoeba* and the choice by inhibitors concentrations was based on previous studies[15–22]. After 1 h, the cells were infected with C. getuliensis at an MOI of 5. Control groups of untreated infected amoebas were also prepared. Thirty minutes post-infection, cells and supernatant were collected and centrifuged at 800 g per 10 minutes. The resultant pellet was washed three times with Page's amoeba saline (PAS)[13]. After, cells were submitted to three rounds of freezing and thawing, to allow the viral particles release, and then subjected to titration using the endpoint method[13,14]. In parallel, the supernatant of cytochalasin assay was also submitted to titration for comparison.

To evaluate the role of cell membranes in the viral replication cycle, 5×10^5 *A. castellanii* cells were also infected with C. getuliensis at an MOI of 5. Thirty minutes post-infection, the amoebas were washed with PAS and then transferred to 6-well microplates containing 1 mL of PYG medium and maintained at 32 °C. After 1 h, brefeldin A (BFA), an inhibitor of membrane traffic, was added at a final concentration of 10 μM, and at 8 and 24 h post-infection, the amoebas were collected for TEM analysis and titration, respectively.

All experiments were performed in triplicate. Graphs were constructed using GraphPad Prism version 7.00 for Windows (GraphPad Software).

Transmission electron microscopy and Scanning electron microscopy. For transmission electron microscopy (TEM), 7×10^6 *Acanthamoeba castellanii* cells were subjected to an asynchronous viral infection using a low MOI of 0.1, and 24 hours post-infection they were recovered and pelleted for 10 min at 800 g. The pellet was washed twice with 0.1 M phosphate buffer (pH 7.4) and fixed with 2.5% glutaraldehyde in 0.1 M phosphate buffer for 1 h at room temperature. The pellet was then washed twice with 0.1 M phosphate buffer and resuspended in the same buffer. After repelleting, the amoebas were embedded in Epon resin by using a standard method, as follows: 2 h of fixation in 2% osmium tetroxide, five washes in distilled water, overnight incubation in uranyl acetate 2% at 2–8 °C, two washes in distilled water, 10 min dehydration in increasing ethanol concentrations (35%, 50%, 70%, 85%, 95% and 100% ethanol), 20 min incubation in acetone and embedding in EPON resin. Ultrathin sections were subsequently analyzed under transmission electron microscopy (TEM; Spirit Biotwin FEI-120 kV).

For scanning electron microscopy assays, 10 μL of purified particles of C. getuliensis 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 same procedure was performed to observe *Acanthamoeba* cell interactions with C. getuliensis during the early (1 h.p.i) and late stages (24 h.p.i.) of infection. The samples were washed three times with 0.1 M cacodylate buffer and post-fixed with 1.0% osmium tetroxide for 1 h at room temperature. After a 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 10 min dehydrated by serial passages in ethanol solutions (35%, 50%, 70%, 85%, 95% and 100%). Samples were subsequently subjected to critical point drying

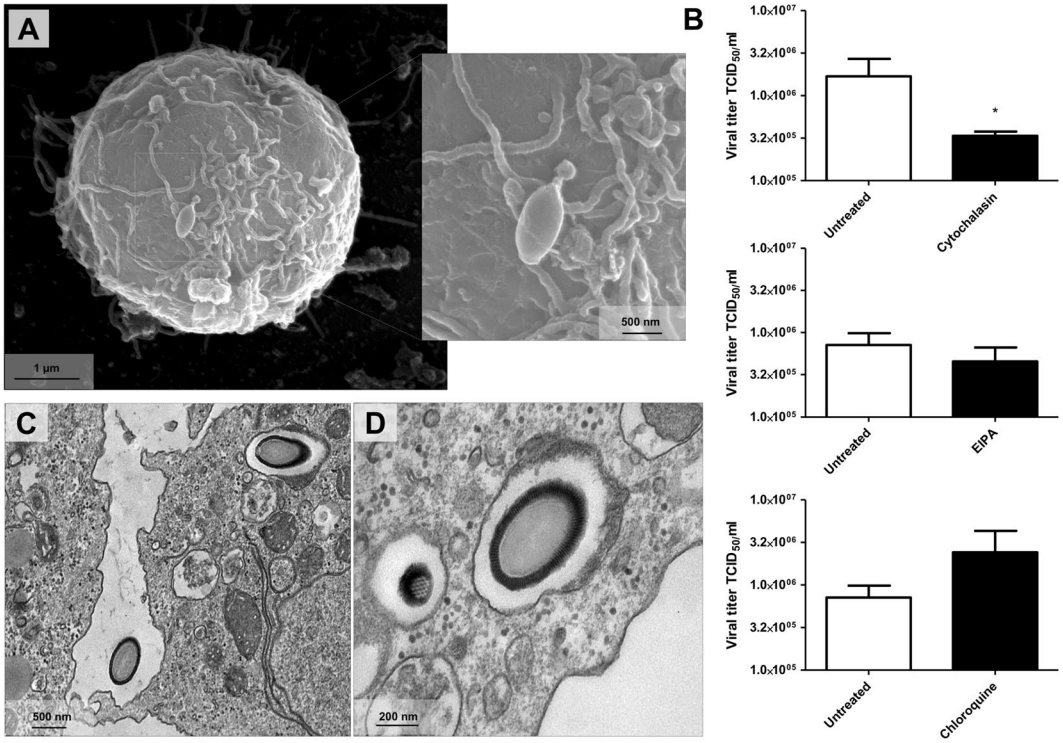


Figure 1. Cedratvirus getuliensis entry in Acanthamoeba castellanii cells. (**A**) Scanning microscopy showing a C. getuliensis particle attached to an *Acanthamoeba* cell. (**B**) The impact of different inhibitors of endocytic pathways in C. getuliensis entry. Treatment of amoebas with cytochalasin D reduced Cedratvirus getuliensis virion incorporation, indicating that particles can enter amoebas by phagocytosis. (**C**) and (**D**) TEM of C. getuliensis particles inside vesicles that strong resemble phagosomes.

using CO_2, placed in stubs and metalized with a 5 nm gold layer. The analyses were completed using scanning electron microscopy (FEG Quanta 200 FEI).

Results

Cytochalasin impacts the incorporation of cedratvirus getuliensis particles by *Acanthamoeba castellanii* cells. Upon discovery of the first cedratviruses[10], analyses involving transmission electron microscopy and scanning electron microscopy[12] revealed the presence of particles with a similar morphology presented by other described viruses, such as pandoravirus and pithovirus. In addition to the morphological similarity, other aspects involving the replication cycle of these viruses were extrapolated and applied to characterize the cedratviruses, such as the internalization of viral particles in amoeba cells by phagocytosis. Our data indicate that Cedratvirus getuliensis can explore phagocytic pathways to enter *A. castellanii* cells, since the titration of pellet cells pretreated with cytochalasin D revealed a significantly decrease (p-value = 0.0385) in the viral titer, when compared to the untreated cells (Fig. 1A and B). Corroborating with those results, when we performed the titration of the supernatant, we observed a higher viral titer for samples pretreated with cytochalasin D, evidencing that a significant number of particles were not phagocytosed (p-value = 0.0243). Transmission electron microscopies of infected particles also corroborate this hypothesis, once C. getuliensis particles could be observed inside vesicles that strong resemble phagosomes (>500 nm), which is consistent with previous studies in which phagocytosis was investigated in amoebas (Fig. 1C and D)[23,24]. In contrast to cytochalasin D, pretreatment with EIPA did not result in a significant reduction in viral titer, indicating that the macropinocytosis is not essential for Cedratvirus getuliensis entry (Fig. 1B). However, as the effects of EIPA has not been previously studied in *Acanthamoeba*, the entry of cedratvirus getuliensis by macropinocytosis cannot be ruled out. In addition, some works have demonstrated that cytochalasin can also interfere on macropinocytosis, that's why a in depth characterization of EIPA in *Acanthamoeba* would be important. Interestingly, we also observed a strong biological tendency of viral title increasing when *Acanthamoeba* cells were treated with chloroquine, an inhibitor of clathrin and caveolin pathways (Fig. 1B).

Cedratvirus getuliensis infection induces the formation of an electron-lucent viral factory and causes cytoplasmic modifications involving different organelles. The replication of many viruses occurs in subcellular microenvironments designated viral factories that originate from the reorganization of cytoskeleton, organelles and cellular membrane compartments[25]. Similarly, the morphogenesis of cedratviruses, as other giant viruses[15,26], occurs in a viral factory located in the cytoplasm of host cells. Using TEM images of Cedratvirus getuliensis replication cycle, we observed the presence of an evident viral factory that in general is as large as the cellular nucleus. Different from mimiviruses, which present an electron-dense viral factory

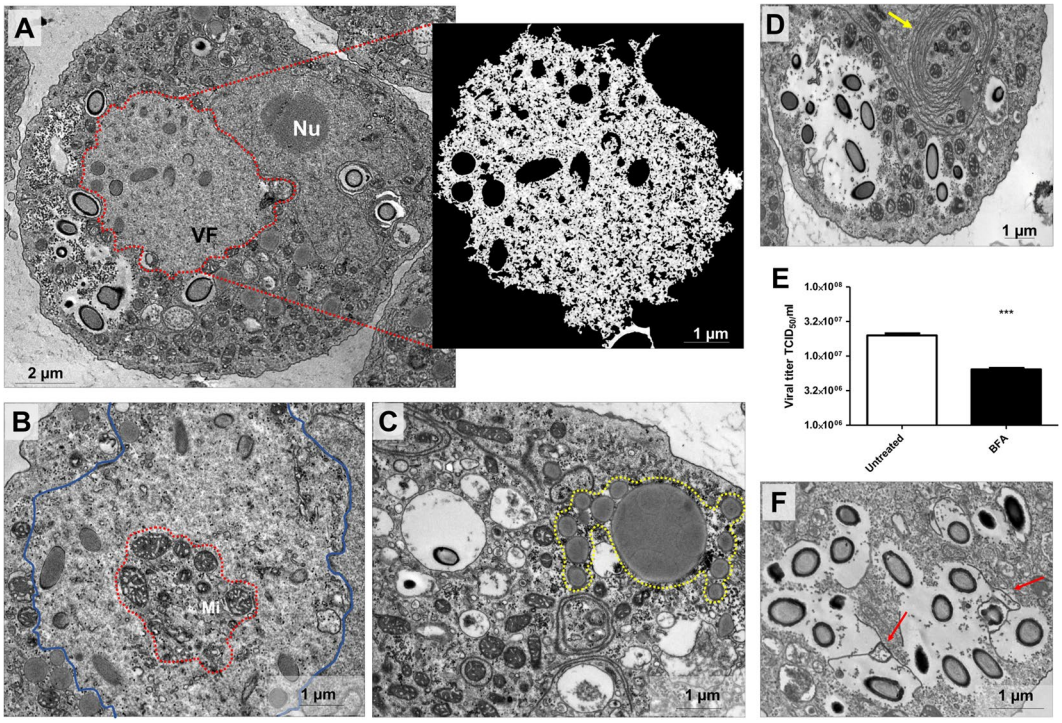


Figure 2. Electron-lucent viral factory and cytoplasmic modifications induced by Cedratvirus getuliensis modification. (**A**) C. getuliensis presents an electron-lucent viral factory (contoured in red and in detail) not easily distinguished from the rest of the cytoplasm and observed at the perinuclear region. Different stages of viral particle morphogenesis could also be observed within the viral factory. (**B**) Abundant presence of mitochondria inside (contoured in red) the viral factory (contoured in blue). (**C**) Lysosomal accumulation and polarization in the host cytoplasm (contoured in yellow). (**D**) Intensified membrane traffic in the host cytoplasm (yellow arrow). (**E**) Treatment with BFA reduces the viral titer after 24 hours of infection. (**F**) Infected cells treated with BFA presented membrane degradation after 8 hours of infection. VF: Viral factory. Nu: Nucleus. Mi: Mitochondria. Image A-right was obtained by TEM and graphically highlighted by using IOS image visualization software.

divided into different parts (one related to genome replication and morphogenesis and another one associated with fibrils acquisition) and are easily distinguished from the rest of the host cytoplasm, the C. getuliensis viral factory is electron-lucent and does not exhibit well-defined zones, thus preventing its prompt distinction from the remaining cytoplasm (Fig. 2A)[26-28]. Moreover, the morphogenesis of C. getuliensis progeny could be observed in the periphery and in the middle of the viral factory, where some electron-dense structures were observed, in contrast to the results observed for mimiviruses, for which the final assembly of new particles occurs at the edge of the factory (Fig. 2A).

Interestingly, we observed that the C. getuliensis viral factory is typically situated at the perinuclear region. During the cycle, the nucleus was consistently present and apparently not affected by the virus, different from that described for pandoraviruses, in which some nuclear disorganization with numerous membrane invaginations were observed in infected cells[11,29]. In addition, during C. getuliensis replication, some absorbing cellular alterations were observed (Fig. 2B,C and D). One alteration was the abundant presence of mitochondria inside and around the viral factory (Fig. 2B). Another interesting change was the intense accumulation and polarization of structures that resemble lysosomal vesicles in the host cytoplasm, particularly during the late steps of the cycle (Fig. 2C). Finally, we also observed exacerbated membrane traffic (Fig. 2D), revealed as important for the morphogenesis and/or exocytosis release of virions, upon the treatment of amoebas with BFA, which significantly impacts the viral titer after 24 hours of infection (Fig. 2E). TEM images also showed a decrease of membrane traffic, as well as membrane degradation in BFA-treated cells, after 8 h of infection (Fig. 2F).

Cedratvirus getuliensis morphogenesis involves the complex and unique sequential organization of immature particles. C. getuliensis morphogenesis is a complex process involving the formation of subsequent structures that could be clearly visualized as electron-dense materials within and at the periphery of the viral factory in TEM images (Fig. 3). TEM images should be analyzed with cautious, since 2D perspective can lead to misinterpretation. However, the obtained images suggest that the first discernible viral particle structures are crescent-shaped ~100 nm precursors developed in the middle of viral factory (Fig. 3A). Similar structures, described as open membrane intermediates or precursors, have been also observed during Vaccinia virus, Mimivirus and African Swine fever virus replication, suggesting the occurrence of a common assembly steps for NCLDVs[30–33]. The following observed differentiation is the longitudinal elongation of the particle (~600 nm),

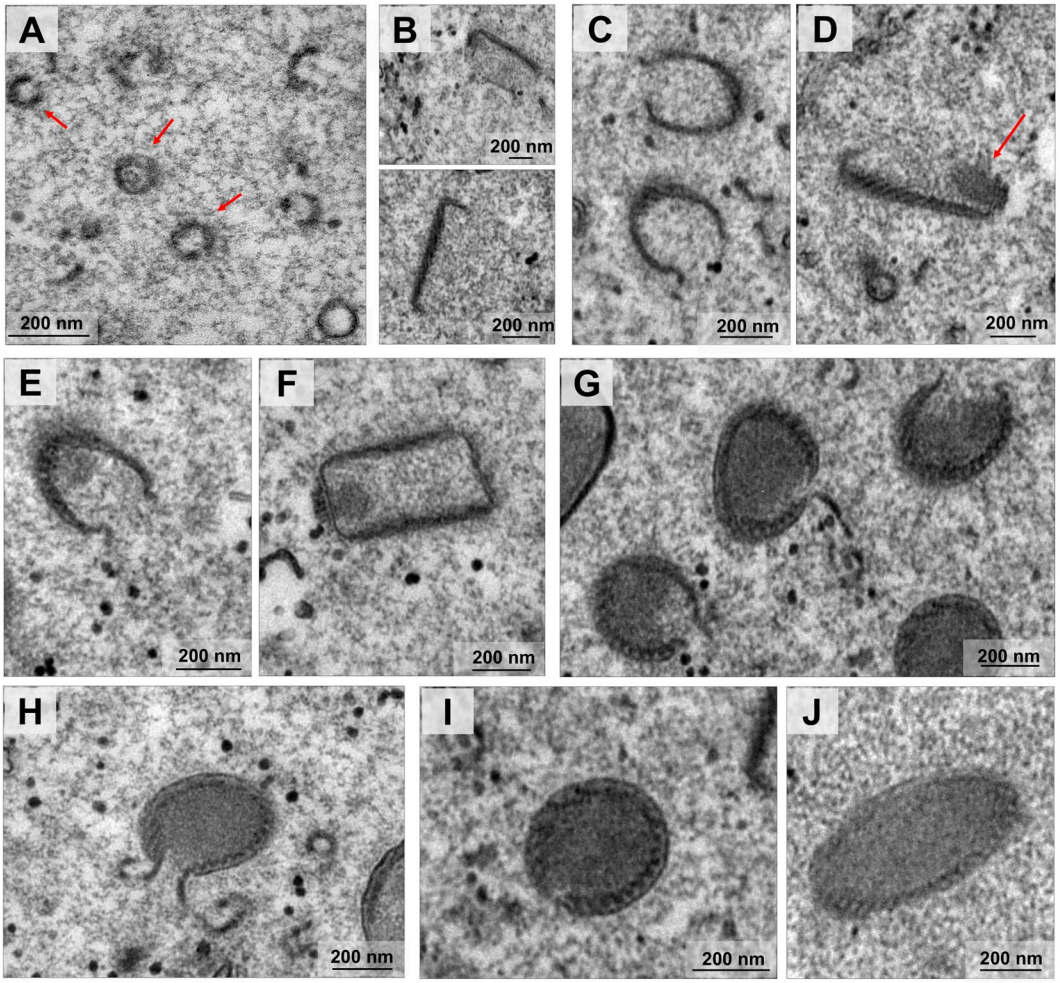


Figure 3. Cedratvirus getuliensis morphogenesis involves the occurrence of subsequent complex structures. (**A**) First discernible viral structures showing a crescent-shaped capsid precursor. (**B**) Longitudinal sections revealed the longitudinal-elongation of the particle and capsids assuming a staple-shaped conformation. (**C**) Transversal sections showed empty capsids with a horseshoe conformation. (**D**) Longitudinal sections showed staple-shaped with the first cork evident (red arrow). The particle may be an open cylinder at this moment, since transversal-sectioned particles appears as rectangles (**F**) and transversal-sectioned particles still reveals horseshoe-like structures (**E**). Progressive filling of the capsid (**G,H**). (**J**) Complete closure of the capsids.

when the precursor capsid assumes a staple-shaped conformation, as visualized by longitudinal sections (Fig. 3B). Transversal sections revealed empty capsids with a similar horseshoe conformation and an evident striated wall, a characteristic feature of pithoviruses as also observed[4,34] (Fig. 3C). At this stage, only one cork region is clearly visible in the longitudinal cut, at the pole where the morphogenesis probably started (Fig. 3D). The particle appears to be an open cylinder at this moment, since longitudinal-sectioned particles appear as rectangles (Fig. 3F) and transversal-sectioned particles still reveal horseshoe-like structures (Fig. 3E). Next, we observed a progressive filling of the capsid (Fig. 3G,H and I), followed by the complete closure of the capsid (Fig. 3J) and the emergence/incorporation of the second cork.

Following the total closure of the capsid, we observed that this structure undergoes some degree of differentiation related to the capsid wall thickness. Immediately after capsid closure, some ovoid particles are observed in the periphery of the viral factory and particle thickening occurs in an area at the edge of or surrounding the viral factory (Fig. 4A and B). Initially, the capsid presents a thin wall and the two corks are not completely laterally covered (Fig. 4C,D and E). As the maturation progresses, the capsid wall becomes thicker until it acquires the same thickness presented by both corks (Fig. 4F,G and H).

Misshapen Cedratvirus getuliensis particles could be observed during virus morphogenesis. We also observed the appearance of some misshapen structures as blobs comprising portions of corks, capsids, membrane and electron-dense material (Fig. 5A and B). These unusual structures have previously been described by Legendre and colleagues in the Pithovirus sibericum replication cycle as "possible reservoirs of partially organized virion building blocks"[4]. We could not discard the hypothesis that these structures might be premature or defective particles, as the occurrence of abnormal particles has previously been reported for other viruses,

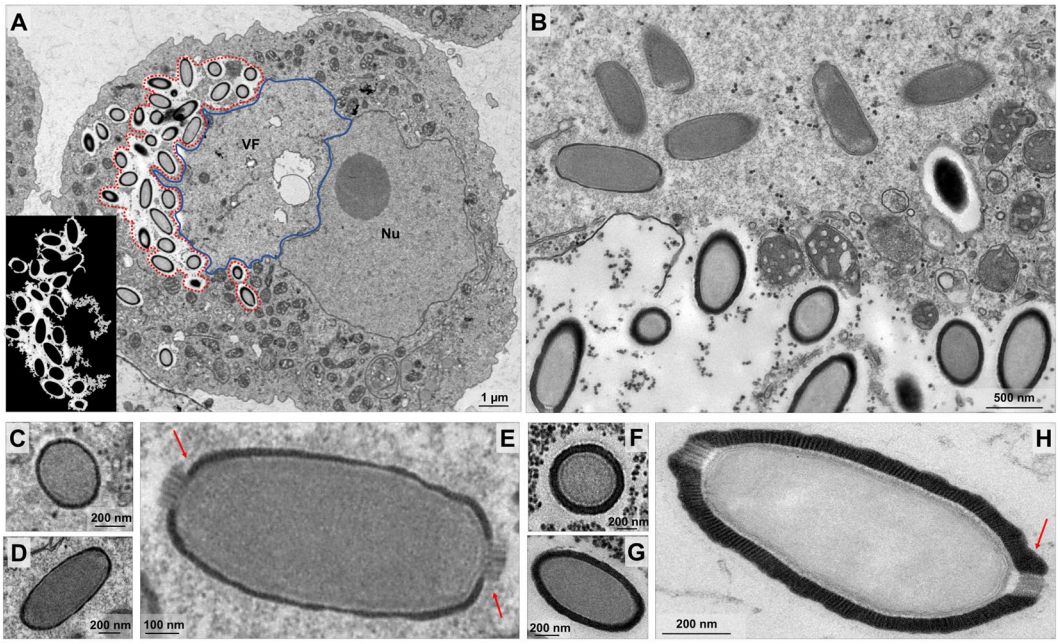


Figure 4. Particle wall thickening after capsids closure. (**A**) Viral particles suffer differentiation related to the thickness in a specific area at the edge of the viral factory (contoured in red and in detail). (**B**) Viral factory periphery evidencing the capsid wall thickness. Cross (**C**) and longitudinal (**D**) sections of capsids presenting a thin thickness and the corks not completely laterally covered (**E**) (red arrows). (**F**) and (**G**) The capsids become thicker with the progression of maturation and acquire the same thickness presented by both corks (**H**) (red arrow). VF: Viral factory. Nu: Nucleus. Image A-left was obtained by TEM and graphically highlighted by using IOS image-visualize software.

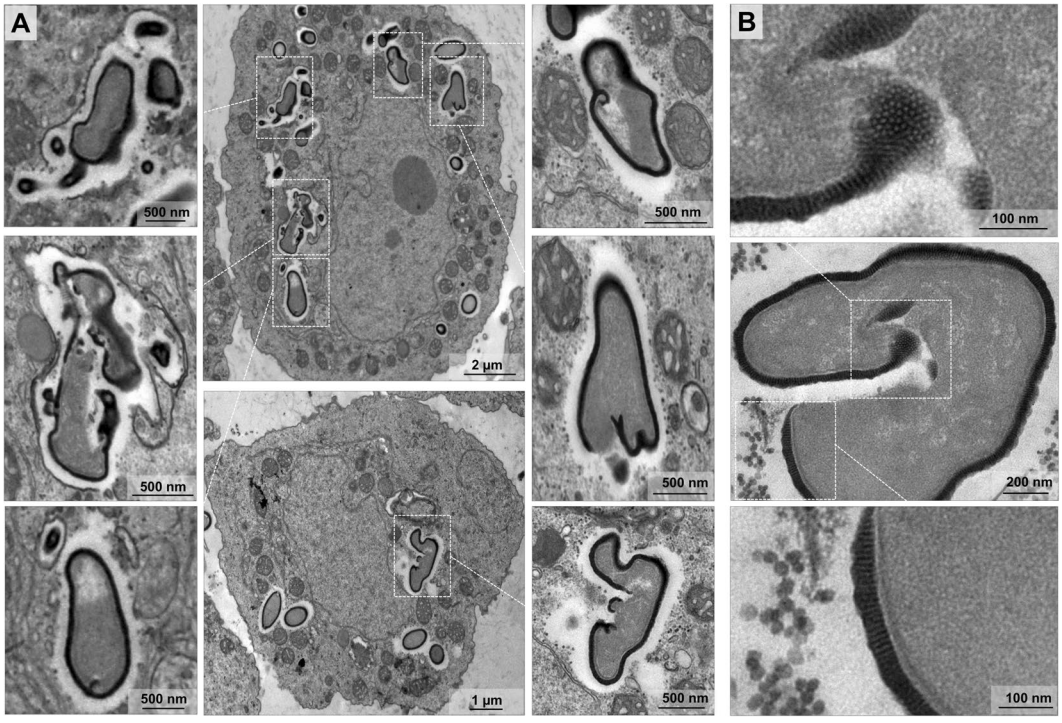


Figure 5. Misshapen structures observed during C. getuliensis multiplication. (**A**) and (**B**) Amorphous structures resembling defective particles and composed by portions of corks, striated capsids, membrane and electron-dense material could be visualized in different regions of the host cytoplasm alongside mature virus.

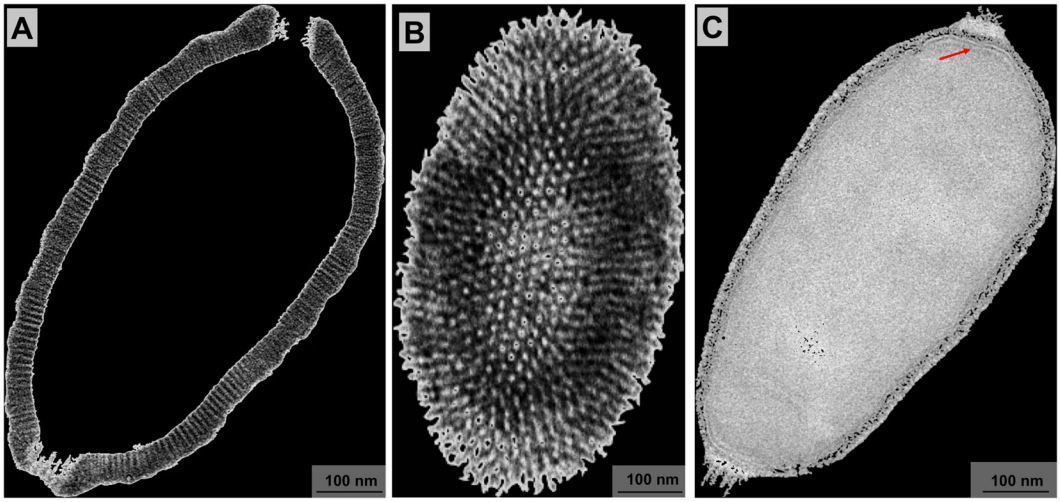


Figure 6. Cedratvirus getuliensis particles present a striped amphora-shaped format and a size plasticity. (**A**) Typical capsid presenting parallel stripes and not completely opposite corks. (**C**) Superficial section of a mature particle evidencing the striped wall. (**D**) Capsid interior composed by a membrane (red arrow) that delimits the internal homogeneous compartment. Images were obtained by TEM and graphically highlighted by using IOS image visualization software.

including giant viruses[26,35]. Notably, the quantification of occurrence of these misshapen structures revealed that 7% of the cells presented at least one of these elements detected in different regions of the host cytoplasm alongside mature virions; thus, these particles are not confined to a viral factory, suggesting that these elements might be defective particles and not particles under morphogenesis.

Viral progeny present ovoid-shaped format, striped capsid and size plasticity. As mentioned, the end of the C. getuliensis replication cycle is characterized by cell lysis with the consequent release of viral particles. An observation of the viral progeny revealed mature particles measuring ~1 µm in size and ~0.5 µm in diameter and showing an ovoid-shaped format with a typical capsid presenting parallel stripes (Fig. 6A). We sagittally sectioned the lateral top, revealing that the virion subunits appear as organized dots (Fig. 6B). Inside this capsid, we observed a putative membrane delimiting the internal compartment without substructures (Fig. 6C). We believe that this putative inner membrane is acquired during the first steps of morphogenesis, prior to the filling of the particles with the viral genome and virion proteins. Unlike that observed for pithovirus, the interior of Cedratvirus getuliensis virions does not harbor episodic electron-dense spheres or tubular structures but is rather homogeneous[4].

As a hallmark of cedratvirus virions, C. getuliensis particles also showed two characteristic protruding striped corks at each apex (Fig. 6A and C). However, although these corks are located at the apices, these structures are not antipodally aligned to each other (Fig. 6A and B) and we observed the existence of a misalignment between the centers of the opposite corks.

Although most of the C. getuliensis particles present a similar morphological pattern, different mature particles were also present. This variation is primarily related to the size of the particles, as shown by scanning electron microscopy analyses that revealed the presence of virions up to 2.04 µm, almost the double the size observed for the majority of particles. Therefore, these data provide evidence of size plasticity for the progeny of Cedratvirus getuliensis, as demonstrated for Pithovirus sibericum[36].

Cedratvirus getuliensis virions can be released after cell lysis or by exocytosis. After the capsid thickening process, the viral morphogenesis and maturation is now complete and new virions are found immersed in the host cytoplasm surrounded by a halo that, despite could be an artefact of epon embedding, is recurrently observed around other giant viruses particles studies[4,6,10,12,37]. Furthermore, new viruses were also observed embedded within membranes (Fig. 7A). Interestingly, these data revealed the presence of one or more particles inside the same vacuole (Fig. 7B,C and D), which could also present more than one membrane (Fig. 7E). We also observed some particles insides vacuoles and outside the cell membrane, but based only in a 2D perspective we could not affirm that the particles are indeed outside of the amoebas or inside some membrane protrusions (Fig. 7F). The presence of the giant virus progeny inside vacuoles has previously been described for Pithovirus sibericum, suggesting that these particles could be released from the cell by exocytosis[4]. Although exocytosis could be an alternative mechanism used for releasing viral progeny, the main strategy used for Cedratvirus getuliensis is cell lysis. Scanning electron microscopy analyses of the late steps of the C. getuliensis cycle reveals the presence of many cells with substantial damage in the plasmatic membrane, where new viral particles are released (Fig. 7G). Furthermore, the cell lysis is accompanied by plasma membrane blebbing (Fig. 7H), that was not visible in control cells not infected by C. getuliensis (Fig. 7I). However, the causes of these blebs induced upon Cedratvirus getuliensis infection deserve further investigation.

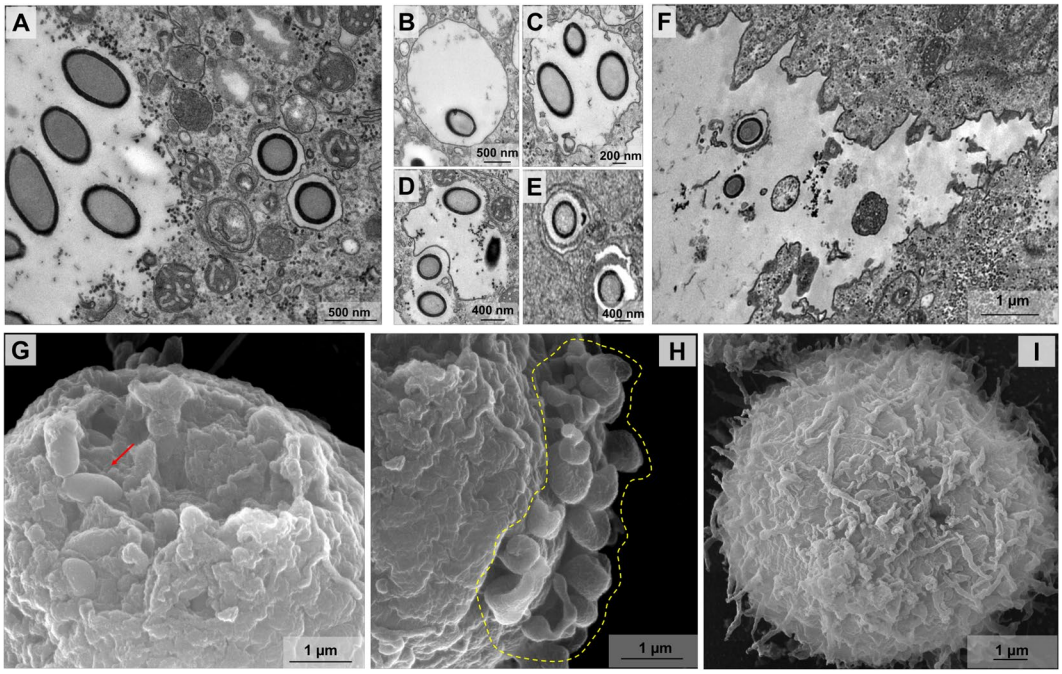


Figure 7. Cedratvirus getuliensis virions can be released after cell lysis or by exocytosis. (**A**) New Cedratvirus getuliensis particles are found immersed in the host cytoplasm or inside vacuoles. (**B–C**) Vacuoles presenting one or more visible particles. (**D**) Particles being engulfed by a membrane. (**E**) Vacuole with more than one membrane. (**F**) Particle insides vacuole apparently outside the cell membrane. (**G**) Scanning microscopy of a host cell presenting a huge damage in the membrane from where new viral particles were released (red arrow). (**H**) Many blebs in the plasma membrane can be observed at the end of infection. (**I**) Cell control not presenting blebs formation in membrane.

Discussion

The current understanding of the virosphere has dramatically changed after the discovery of mimivirus, which paved the way for the discovery of other giant and complex amoeba-infecting viruses[29]. Although many studies have highlighted that giant viruses can be phylogenetically related and may form a new putative viral order 'Megavirales' along with other large DNA viruses[38], these viruses present a plethora of virion structures and remarkable differences regarding their developmental cycles. In the present study, we present the first in-depth description of Cedratvirus getuliensis replication cycle, providing valuable information to better understand the biology of this new group of viruses.

Cedratviruses are ~1.4 µm in size and ~0.5 µm in diameter, representing one of the longest viruses described thus far, along with their close relative pithoviruses[4,10,12,34]. Due to their huge size, it was initially proposed that these viruses started their replicative cycle by entering the hosts through phagocytosis, but no experimental data was provided to support this hypothesis, except for a few microscopy images. Here, we demonstrated that the inhibition of phagocytosis with cytochalasin D results in a reduction of Cedratvirus getuliensis virion incorporation by amoebas, suggesting that this virus may enter by this pathway (Fig. 1B,C and D). However, the inhibition of macropinocytosis by EIPA does not affect the entry of Cedratvirus getuliensis particles (Fig. 1B). Interestingly, *Acanthamoeba* cells treatment with chloroquine increased C. getuliensis viral titer, suggesting that this inhibitor could accumulates inside the phagosomes, resulting in pH increasing and consequent prevention of uncoating process; thus preserving a higher number of not uncoated virions inside phagosomes. Following entry, one of the corks is expelled, enabling the fusion of the internal membrane with the phagosome membrane and further releasing the genome into the host cytoplasm[10]. The precise mechanism that triggers these events remains unclear, but it may be related to the low pH environment of phagosomes, similar to the mechanism observed for mimiviruses[26].

After an eclipse phase, a large electron-lucent viral factory (VF) is formed, wherein genome replication and virion morphogenesis occur. It is still uncertain whether the host nucleus is involved in the replication of the cedratvirus genome, since the nucleus remains apparently unaltered during the entire viral cycle, different from other giant viruses[6,11]. Similar to pithovirus, no delimiting structure was observed around the VF of Cedratvirus getuliensis, which is perinuclearly located[4]. Cedratviruses present a gene-set related to DNA replication and transcription[10,12], and it is possible that these elements are packaged into mature virions, similar to its closest relative Pithovirus sibericum[4]; no nuclear machinery is required during cedratvirus replication, in contrast to that described for marseillevirus[39]. The morphogenesis of cedratviruses is complex, wherein different structures are observed until the full maturation of the virion, which exclusively occurs within the VF (Figs. 3 and 4). Similar to other large and giant DNA viruses, cedratviruses form crescent-like structures and may exhibit an internal membrane, although its origin is still unknown[15,26,30–33,39]. Besides to this putative intern membrane, we also observed transversal-sectioned capsids been filled with an electron-dense material that suggest the occurrence of genome and virion protein acquisition (Fig. 3G,H and I). The complete morphogenesis of the virion resembles that of

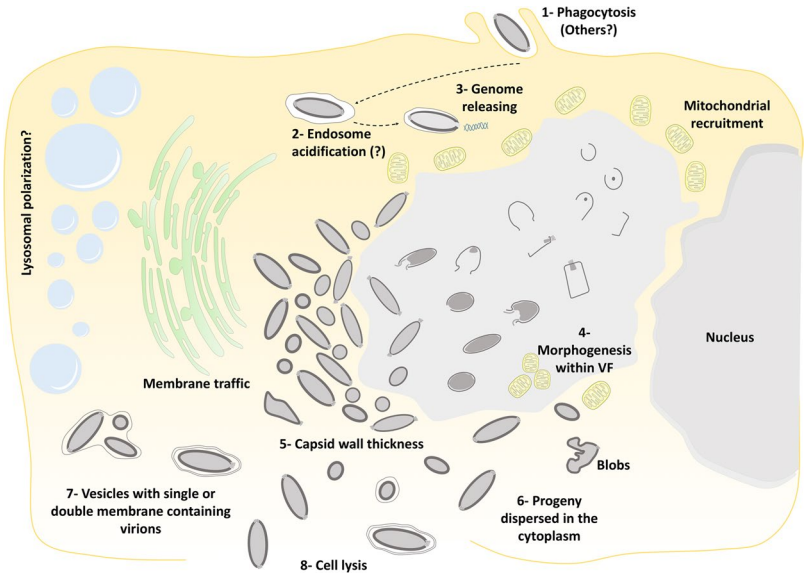


Figure 8. Schematic representation of the Cedratvirus getuliensis replication cycle. Infectious particles enter their host cells by phagocytosis (1). Following entry, one cork is expelled and the fusion of the internal membrane with the phagosome membrane occurs with the release of the genome into the cytoplasm, as demonstrated by Andreani *et al.* (2016) and Bertelli *et al.* (2017) (2–3). The precise mechanism that triggers the cork expelling and membrane fusion remains unknown, but it may be related to the low pH environment of phagosome. After an eclipse phase, a large electron-lucent viral factory is formed, wherein the virion morphogenesis occurs (4). This process is complex and involves sequential structures starting with crescent-shaped precursors, followed by longitudinal-elongation, first cork acquisition, and the presence of structures with a horseshoe-like and rectangular shape (4). Next, a progressive filling of the capsid is observed, followed by the complete closure of the capsid and the emergence of the second cork (4). Thickening of the capsid wall (5) also occurs, and new viral particles are observed dispersed throughout the cytoplasm (6) or taken up by vesicles with one or two membranes (7). The viral progeny are primarily released through cell lysis (8), but exocytosis is also likely to occur.

pithoviruses, with a rectangular shape initially emerging, followed by a thickening of the capsid and subsequent acquisition of an oval shape[4]; but differently from its relative, cedratviruses acquire a second cork at the end of the process. Furthermore, no horseshoe structure has been described for pithoviruses. It is likely that this feature is shared by the members of the putative '*Pithoviridae*' family, but additional studies on the morphogenesis of pithoviruses are needed to corroborate this hypothesis. The replication cycle is completed with the release of new viral particles primarily through cell lysis, but exocytosis is likely to occur, since we observed some viral particles embedded in the membranes and outside the host cells. The origin of these membranes is not clear, but we observed that treatment with BFA significantly impacted the viral titer, showing that membrane traffic is important for the occurrence of virion morphogenesis and/or exocytosis. Although no specific labeling for lysosomes was used, we observed the polarization of structures that resemble these organelles during cedratviruses infection, that could suggest the occurrence of autophagy of target viral components or virions, once this organelle acts as an end point degradative structure (Fig. 2)[40]. Moreover we also noted the recruitment of mitochondria, which could be related to the optimization of energy acquisition, required for viral replication (Fig. 2)[40]. However, the actual impact of these organelles on the viral replication cycle remains unknown. Finally, based on the present data, we provide a general view of the entire life cycle of cedratviruses (Fig. 8 – see legend for details).

There are still some unanswered questions concerning the replication cycle of this new group of viruses, especially at the molecular level. Further investigation using different imaging techniques, combined with transcriptomics and proteomics data, will certainly provide valuable insights into the virus-host interaction dynamics and fill some remaining gaps concerning the life cycle of cedratviruses. The world of giant viruses is constantly increasing, and investigating their infectious biology will provide a better understanding of the ecology and evolution of these complex organisms.

References

1. La Scola, B. *et al.* A giant virus in amoebae. *Science* **299**, 2033, https://doi.org/10.1126/science.1081867 (2003).
2. Pagnier, I. *et al.* A decade of improvements in Mimiviridae and Marseilleviridae isolation from amoeba. *Intervirology* **56**, 354–363, https://doi.org/10.1159/000354556 (2013).
3. Dornas, F. P. *et al.* Isolation of new Brazilian giant viruses from environmental samples using a panel of protozoa. *Frontiers in microbiology* **6**, 1086, https://doi.org/10.3389/fmicb.2015.01086 (2015).
4. 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**, 4274–4279, https://doi.org/10.1073/pnas.1320670111 (2014).

5. Saadi, H. *et al*. First isolation of Mimivirus in a patient with pneumonia. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **57**, e127–134, https://doi.org/10.1093/cid/cit354 (2013).
6. 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**, E5327–5335, https://doi.org/10.1073/pnas.1510795112 (2015).
7. Lwoff, A. Interaction among virus, cell, and organism. *Science* **152**, 1216–1220 (1966).
8. Bajrai, L. H. *et al*. Kaumoebavirus, a New Virus That Clusters with Faustoviruses and Asfarviridae. *Viruses* **8**, https://doi.org/10.3390/v8110278 (2016).
9. Reteno, D. G. *et al*. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *Journal of virology* **89**, 6585–6594, https://doi.org/10.1128/JVI.00115-15 (2015).
10. Andreani, J. *et al*. Cedratvirus, a Double-Cork Structured Giant Virus, is a Distant Relative of Pithoviruses. *Viruses* **8**, https://doi.org/10.3390/v8110300 (2016).
11. Philippe, N. *et al*. 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 (2013).
12. Bertelli, C. *et al*. Cedratvirus lausannensis - digging into Pithoviridae diversity. *Environmental microbiology* **19**, 4022–4034, https://doi.org/10.1111/1462-2920.13813 (2017).
13. Abrahao, J. S. *et al*. Mimiviruses: Replication, Purification, and Quantification. *Current protocols in microbiology.* **41**(14G), 11 11–14G 11 13, https://doi.org/10.1002/cpmc.2 (2016).
14. Reed, L. J. M. H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg* **27**, 493–497 (1938).
15. Arantes, T. S. *et al*. The Large Marseillevirus Explores Different Entry Pathways by Forming Giant Infectious Vesicles. *Journal of virology* **90**, 5246–5255, https://doi.org/10.1128/JVI.00177-16 (2016).
16. Moon, E. K. *et al*. Autophagy inhibitors as a potential antiamoebic treatment for Acanthamoeba keratitis. *Antimicrobial agents and chemotherapy* **59**, 4020–4025, https://doi.org/10.1128/AAC.05165-14 (2015).
17. Jha, B. K. *et al*. Chloroquine has a cytotoxic effect on Acanthamoeba encystation through modulation of autophagy. *Antimicrobial agents and chemotherapy* **58**, 6235–6241, https://doi.org/10.1128/AAC.03164-14 (2014).
18. Ghigo, E. *et al*. Ameobal pathogen mimivirus infects macrophages through phagocytosis. *PLoS pathogens* **4**, e1000087, https://doi.org/10.1371/journal.ppat.1000087 (2008).
19. Alsam, S., Sissons, J., Dudley, R. & Khan, N. A. Mechanisms associated with Acanthamoeba castellanii (T4) phagocytosis. *Parasitology research* **96**, 402–409, https://doi.org/10.1007/s00436-005-1401-z (2005).
20. Chrisman, C. J., Alvarez, M. & Casadevall, A. Phagocytosis of Cryptococcus neoformans by, and nonlytic exocytosis from, Acanthamoeba castellanii. *Applied and environmental microbiology* **76**, 6056–6062, https://doi.org/10.1128/AEM.00812-10 (2010).
21. Soto-Arredondo, K. J., Flores-Villavicencio, L. L., Serrano-Luna, J. J., Shibayama, M. & Sabanero-Lopez, M. Biochemical and cellular mechanisms regulating Acanthamoeba castellanii adherence to host cells. *Parasitology* **141**, 531–541, https://doi.org/10.1017/S0031182013001923 (2014).
22. Taylor, W. M., Pidherney, M. S., Alizadeh, H. & Niederkorn, J. Y. *In vitro* characterization of Acanthamoeba castellanii cytopathic effect. *The Journal of parasitology* **81**, 603–609 (1995).
23. Korn, E. D. & Weisman, R. A. Phagocytosis of latex beads by Acanthamoeba. II. Electron microscopic study of the initial events. *The Journal of cell biology* **34**, 219–227 (1967).
24. Wetzel, M. G. & Korn, E. D. Phagocytosis of latex beads by Acahamoeba castellanii (Neff). 3. Isolation of the phagocytic vesicles and their membranes. *The Journal of cell biology* **43**, 90–104 (1969).
25. de Castro, I. F., Volonte, L. & Risco, C. Virus factories: biogenesis and structural design. *Cellular microbiology* **15**, 24–34, https://doi.org/10.1111/cmi.12029 (2013).
26. Andrade, A. *et al*. Filling Knowledge Gaps for Mimivirus Entry, Uncoating, and Morphogenesis. *Journal of virology* **91**, https://doi.org/10.1128/JVI.01335-17 (2017).
27. 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, https://doi.org/10.1371/journal.pone.0000328 (2007).
28. Kuznetsov, Y. G., Klose, T., Rossmann, M. & McPherson, A. Morphogenesis of mimivirus and its viral factories: an atomic force microscopy study of infected cells. *Journal of virology* **87**, 11200–11213, https://doi.org/10.1128/JVI.01372-13 (2013).
29. Colson, P., La Scola, B. & Raoult, D. Giant Viruses of Amoebae: A Journey Through Innovative Research and Paradigm Changes. *Annual review of virology* **4**, 61–85, https://doi.org/10.1146/annurev-virology-101416-041816 (2017).
30. Suarez, C. *et al*. African swine fever virus assembles a single membrane derived from rupture of the endoplasmic reticulum. *Cellular microbiology* **17**, 1683–1698, https://doi.org/10.1111/cmi.12468 (2015).
31. Suarez, C. *et al*. Open membranes are the precursors for assembly of large DNA viruses. *Cellular microbiology* **15**, 1883–1895, https://doi.org/10.1111/cmi.12156 (2013).
32. Mutsafi, Y., Shimoni, E., Shimon, A. & Minsky, A. Membrane assembly during the infection cycle of the giant Mimivirus. *PLoS pathogens* **9**, e1003367, https://doi.org/10.1371/journal.ppat.1003367 (2013).
33. Mutsafi, Y., Zauberman, N., Sabanay, I. & Minsky, A. Vaccinia-like cytoplasmic replication of the giant Mimivirus. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 5978–5982, https://doi.org/10.1073/pnas.0912737107 (2010).
34. Levasseur, A. *et al*. Comparison of a Modern and Fossil Pithovirus Reveals Its Genetic Conservation and Evolution. *Genome biology and evolution* **8**, 2333–2339, https://doi.org/10.1093/gbe/evw153 (2016).
35. Abrahao, J. S. *et al*. Acanthamoeba polyphaga mimivirus and other giant viruses: an open field to outstanding discoveries. *Virology journal* **11**, 120, https://doi.org/10.1186/1743-422X-11-120 (2014).
36. Okamoto, K. *et al*. Structural variability and complexity of the giant Pithovirus sibericum particle revealed by high-voltage electron cryo-tomography and energy-filtered electron cryo-microscopy. *Scientific reports* **7**, 13291, https://doi.org/10.1038/s41598-017-13390-4 (2017).
37. Campos, R. K. *et al*. Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. *Virology journal* **11**, 95, https://doi.org/10.1186/1743-422X-11-95 (2014).
38. Colson, P., de Lamballerie, X., Fournous, G. & Raoult, D. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. *Intervirology* **55**, 321–332, https://doi.org/10.1159/000336562 (2012).
39. Fabre, E. *et al*. Noumeavirus replication relies on a transient remote control of the host nucleus. *Nature communications* **8**, 15087, https://doi.org/10.1038/ncomms15087 (2017).
40. Novoa, R. R. *et al*. Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biology of the cell* **97**, 147–172, https://doi.org/10.1042/BC20040058 (2005).

Acknowledgements

The authors would like to thank our colleagues from Gepvig and the Laboratório de Vírus for their excellent technical support. The authors would also like to thank the CNPq, CAPES and FAPEMIG for scholarships and the Center of Microscopy of UFMG. J.S.A., C.A.B. and E.G.K. are CNPq researchers.

Author Contributions

L.K.S.S., A.C.S.P.A., F.P.D., R.A.L.R., T.A.: performed experiments; J.S.A., C.A.B. and E.G.K.: designed the study. L.K.S.S.: made the Fig. 8; L.K.S.S., R.A.L.R., J.S.A.: wrote the manuscript. All the authors read the final version of the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

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

viruses

MDPI

Article

A Model to Detect Autochthonous Group 1 and 2 Brazilian *Vaccinia virus* Coinfections: Development of a qPCR Tool for Diagnosis and Pathogenesis Studies

Rafael Calixto [1], Graziele Oliveira [1], Maurício Lima [1], Ana Cláudia Andrade [1], Giliane de Souza Trindade [1], Danilo Bretas de Oliveira [1,2] and Erna Geessien Kroon [1,*]

[1] Laboratório de Vírus, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Minas Gerais, Brazil; calixtomicro@yahoo.com.br (R.C.); graziufmg@yahoo.com.br (G.O.); maurili15@hotmail.com (M.L.); ana.andrade2008@hotmail.com (A.C.A.); gitrindade@yahoo.com.br (G.d.S.T.); danilobretas@yahoo.com.br (D.B.d.O.)

[2] Faculdade de Medicina de Diamantina, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Dimantina 39100-000, Minas Gerais, Brazil

* Correspondence: kroone@icb.ufmg.br or ernagkroon@gmail.com; Tel.: +55-031-99949-1240

Received: 1 December 2017; Accepted: 29 December 2017; Published: 30 December 2017

Abstract: *Vaccinia virus* (*VACV*) is the etiological agent of bovine vaccinia (BV), an emerging zoonosis that has been associated with economic losses and social effects. Despite increasing reports of BV outbreaks in Brazil, little is known about the biological interactions of Brazilian *VACV* (*VACV*-BR) isolates during coinfections; furthermore, there are no tools for the diagnosis of these coinfections. In this study, a tool to co-detect two variants of *VACV* was developed to provide new information regarding the pathogenesis, virulence profile, and viral spread during coinfection with *VACV*-BR isolates. To test the quantitative polymerase chain reactions (qPCR) tool, groups of BALB/c mice were intranasally monoinfected with Pelotas virus 1—Group II (PV1-GII) and Pelotas virus 2—Group I (PV2-GI), or were coinfected with PV1-GII and PV2-GI. Clinical signs of the mice were evaluated and the viral load in lung and spleen were detected using simultaneous polymerase chain reactions (PCR) targeting the *A56R* (*hemagglutinin*) gene of *VACV*. The results showed that qPCR for the quantification of viral load in coinfection was efficient and highly sensitive. Coinfected mice presented more severe disease and a higher frequency of *VACV* detection in lung and spleen, when compared to monoinfected groups. This study is the first description of PV1 and PV2 pathogenicity during coinfection in mice, and provides a new method to detect *VACV*-BR coinfections.

Keywords: *Vaccinia virus*; qPCR; coinfection; mice model

1. Introduction

Vaccinia virus (*VACV*) is a member of the family *Poxviridae*, genus *Orthopoxvirus*, which includes other members, such as *Variola virus*, *Cowpox virus* and *Monkeypox virus* [1,2]. *Variola virus* was one of the most terrible pathogens in human history, but it was declared eradicated in 1980 after an intensive vaccination campaign promoted by the World Health Organization (WHO) [3]. *VACV* can induce serological cross-reactivity against other orthopoxvirus (OPV) members and was used in the WHO campaign [1,3].

VACV is the etiological agent of bovine vaccinia (BV), an exanthematous disease that causes ulcerative lesions in cattle and humans, economic losses, and social effects in South America and Asia, especially in Brazil [4–8]. The clinical signs of BV range from papules and vesicles to scabs, mainly on the udder and teats of bovines. In humans, lesions occur primarily on the hands and arms, and other

symptoms, such as fever, myalgia, headache, arthralgia, and lymphadenopathy, have been described and there is a significant economic impact on rural workers [1,6].

The natural circulation of *VACV* in Brazil has been often reported since 1999, and is associated with exanthematous outbreaks [6,7,9–12]. Many studies have shown biological and genetic variations among Brazilian *VACV* (*VACV*-BR) isolates. This variability allowed *VACV*-BR clustering into two distinct groups: Group 1 (GI) and group 2 (GII). These two groups are supported by biological features, such as virulence in a BALB/c mouse model and plaque phenotype in BSC-40 cells. GII isolates display larger plaque sizes and are virulent to mice, unlike GI [4,12–16]. Furthermore, molecular diversity is observed in specific *VACV* genes, such as the *hemagglutinin* gene (*A56R*), *A-type inclusion body* gene (*A26L*), and *chemokine-binding protein* gene (*C23L*), and these genes have been used in phylogenetic studies and further confirmed the dichotomy between GI and GII *VACV*-BR. The *A56R* sequence contains a signature deletion of 18 nt, present in the sequences of GI isolates and absent of GII isolates, which is used as a "molecular marker" for *VACV*-BR group identification [9,13,16–18]. The circulation of the two *VACV-BR* groups was demonstrated in the same outbreak in 2006 [4] and in the same host as a coinfection was only identified later [15,16]. In 2008, a *VACV* outbreak, caused by viruses of the two *VACV-BR* groups, was described in horses from Pelotas City, Brazil. This coinfection presented hemorrhagic lesion and scabs in the muzzles and nostrils of animals [15,19]. The *VACV* isolates were named Pelotas virus 1—Group II (PV1-GII) and Pelotas virus 2—Group I (PV2-GI) and were used in this study [15]. Despite increasing reports of outbreaks related to *VACV*-BR, little is known about its biological relevance, virulence profile and viral spread during coinfections with *VACV-BR* of GI and GII. Moreover, until now, there have been no established tools for the diagnosis or pathogenesis studies of coinfections with *VACV-BR* of GI and GII. As the best well-characterized molecular difference of the two groups is the *A56R* sequence which contains a signature deletion of 18 nt for GI, it results in the difficulty of developing a test for direct quantification of this group.

In this study, a new tool for the detection and quantification of *VACV* isolates in coinfections was developed and could be used as a tool to provided new information regarding the diagnosis, pathogenicity, virulence profile, and viral spread during a coinfection with *VACV*-BR isolates. Our method aims to improve screening in outbreaks and consequently the study of *VACV*-BR GI/GII coinfections pathogenesis.

2. Materials and Methods

2.1. Ethical Statement

This study was approved by the Committee of Ethics in Animal Use from the Universidade Federal de Minas Gerais (CEUA/UFMG, Belo Horizonte, Brazil), protocol number 207/2010.

2.2. Cells and Viruses

African green monkey kidney BSC-40 (ATCC-CRL-2761) and VERO (ATCC-CCL-81) cells were maintained in a 5% CO_2 atmosphere at 37 °C, in Eagle's Minimum Essential Medium (MEM) (Gibco BRL, Invitrogen, Carlsbad, CA, USA), supplemented with 5% fetal bovine serum (FBS) (Cultilab, Brazil), 25 µg/mL fungizone (Amphotericin B) (Cristália, São Paulo, Brazil), 500 U/mL penicillin and 50 µg/mL gentamicin (Schering-Plough, São Paulo, Brazil). VERO cells were used for viral replication and the BSC-40 cells were used for viral plaque phenotypes and titration. The *VACV* used in the study, PV1 and PV2, were isolated from clinical specimens of horses during an equine vaccinia outbreak [15,19].

2.3. Animal Experiments

For all animal experiments, five-week-old male Balb/c mice were used, and were maintained in micro-isolators located in a ventilated animal caging system (Alesco Ltd., Campinas, SP, Brazil), and were provided with commercial mouse food and water, ad libitum, in controlled lighting (12 h

light–12 h dark), humidity (60–80%) and temperature (22 ± 1 °C). The mice of all groups were anesthetized via intraperitoneal injection of 0.1 mg of ketamine and 0.01 mg of xylazine in 0.9% PBS, per gram of animal weight. The four groups of five mice were inoculated intranasally with 10 µL of viral suspension: PV1-GII 1×10^6 p.f.u.; PV2-GI 1×10^6 p.f.u.; PV1-GII+ PV2-GI 5×10^5 p.f.u. of each sample; a negative control group was inoculated with 10 µL of PBS, as previously described [14]. Mice were weighed daily, and other clinical signs were recorded for 30 days post infection (d.p.i.). To study viral tropism, the mice were euthanized with an overdose of anesthetics (three times the anesthetic solution) and were perfused with PBS-EDTA intracardiac and those animals had their spleens and lungs collected on 5 d.p.i. [14]. The collected organs were weighed and macerated with the Beadbeater 16 homogenizer in 500 µL of PBS, with glass beads in the microtubes thread. Three cycles of freezing and thawing were performed in order to release the viral particles from the cells. The cells were then centrifuged at 425× *g* for 10 min at 4 °C and the supernatants were used for DNA extraction and plaque phenotype assays. For DNA extraction, the High Pure Viral Nucleic Acid Kit (Roche Diagnostics GmbH, Branchburg, NJ, USA) was used.

2.4. Plaque Phenotype

For plaque phenotype assays, BSC40 cells seeded in 6-well plates at 90–95% confluence were inoculated with the macerated tissues. After 1 h of adsorption (37 °C, 5% CO_2), monolayers were washed twice with PBS and overlaid with solid medium, prepared by mixing equal parts (1:2) of 1% agarose and twice the standard concentration Eagle's minimum essential medium (MEM) supplemented with 2% FBS. After 48 h of incubation (37 °C, 5% CO_2), cells were fixed with formaldehyde and stained with crystal violet for plaque size analysis.

2.5. qPCR

The qPCR consisted of two simultaneous reactions, A and B, targeting the *hemagglutinin* gene [20]. DNA amplifications were carried out in duplicate in a StepOne™ thermocycler (Applied Biosystems, Foster City, CA, USA). qPCR took place in a total volume of 10 µL, containing 5 µL of Master Mix SYBR Green (Applied Biosystems, manufactured by Roche, Branchburg, NJ, USA), 200 nM of each of the forward and reverse primers (Table 1) and 10–50 ng of each DNA sample.

Table 1. Real-time PCR primers.

Reaction	Primers	Sequence (5′-3′)	Specificity
A	A56R-gen F A56R-gen R	AACCACCGATGATGCGGAT TGCCACGGCCGACAATATAA	Amplify all VACV Group I and II.
B	A56R-BVV-nDEL F A56R-generic R	GCGGATCTTTATGATACGTACAATG ACGGCCGACAATATAATTAATGC	Amplify all VACV that do not present the 18nt deletion Group II [20].

Figure 1 shows the target DNA regions for the primers used in this study. Thermal cycling conditions were as follows: one cycle at 95 °C for 10 min, followed by 40 cycles at 95 °C for 10 s, and 58 °C for 40 s; and a melting curve analysis, consisting of 95 °C for 15 s, 58 °C for 15 s, followed by increasing the temperature by 1 °C every 2 s until 95 °C was reached, and then 95 °C for 15 s. Standard curves were constructed by plotting four dilutions of each prototype DNA against the corresponding cycle threshold value (Ct); melting curves were used to ensure there were no non-specific amplifications.

Figure 1. Alignment of the target DNA region within the *A56R* gene of the *Vaccinia virus* used for primer sequence design. The virus sequences were obtained from GenBank and the accession numbers are shown in the figure. The alignment was performed using the standard parameters of CLUSTAL W. The primers used in Reaction A (*A56R*-gen F and *A56R*-gen R) are outlined in black and, for Reaction B (*A56R*-BVV-nDEL F and *A56R*-generic R), are in red.

For standardization, a curve was made using lungs from uninfected Balb/c mice, spiked with 10^6, 10^5, 10^4, 10^3, 10^2 or 10^1 p.f.u. of PV1-GII and PV2-GI. For validation lungs from uninfected Balb/c mice were spiked with 10^5 p.f.u. of each sample (PV1-GII and PV2-GI). DNA was extracted and qPCR reactions A and B were performed. An equation for the differential viral load calculation was proposed:

[ ] *VACV* GI = [ ] *VACV* total (GI+GII) − [ ] *VACV* GII
[ ] *VACV* GI = $3.38 \times 10^6 - 1.68 \times 10^6$
[ ] *VACV* GI = 1.70×10^6
GI = 1.70×10^6 and GII = 1.68×10^6

2.6. Statistical Analyses

All results were plotted using GraphPad Prism (GraphPad Software, version 6.01, La Jolla, CA, USA) and compared using one-way ANOVA and Tukey's multiple comparisons test, and two-way ANOVA using the Bonferroni method. In all tests, p-values < 0.05 were considered statistically significant.

3. Results

3.1. Development of qPCR Tool

The efficiency of primers developed for the *A56R* gene in *VACV* coinfection, for Reaction A, showed an efficiency of 96.6% and an R^2 value of 0.971; for Reaction B, an efficiency of 94.0% and a value of R^2 of 0.983 (Figure 2) were found. The detection limit of the assay was 2 p.f.u./mg of tissue and the reaction was efficient in different matrices such as lung, gut, and spleen.

The analysis of only one quantitative PCR (qPCR) peak fluorescence reaction was not able to distinguish the *VACV*-BR groups and to identify the presence of the deletion of 18 nt in *A56R* gene of GI viruses. In qPCR Reaction A, the fluorescence peaks were within 0.5 °C, making it difficult to rely on the melting (Tm) curve as a unique identifier for each PCR product (Tm GI: 75.8 °C and Tm GII: 76.3 °C). In Reaction B, GI is not amplified.

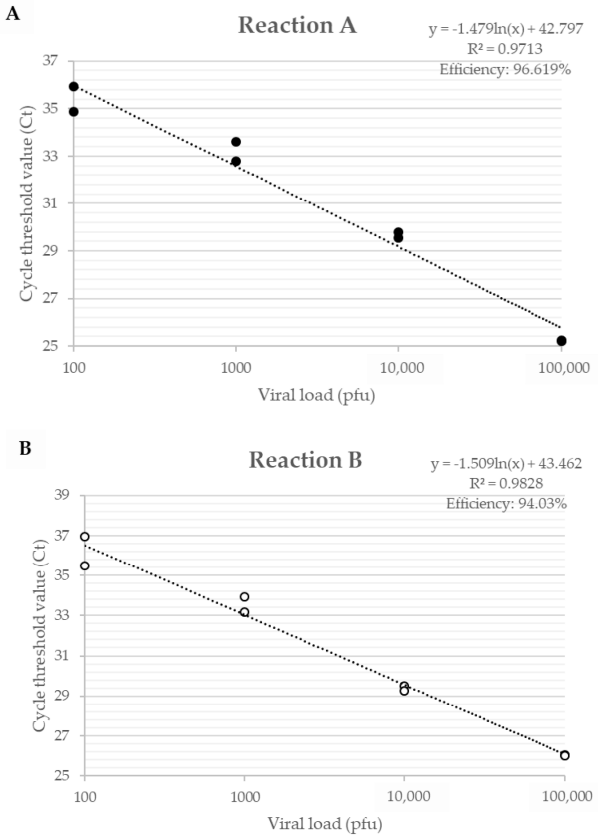


Figure 2. Efficiency curve of *A56R* gene qPCR in coinfection of PV1-GII and PV2-GI. Uninfected Balb/c mice lungs were spiked with 10^6, 10^5, 10^4, 10^3, 10^2 and 10^1 p.f.u. of PV1-GII and PV2-GI. (**A**) in **Reaction A shown in black circles**, primers *A56R*-gen F and *A56R*-gen R were used; and (**B**) in **Reaction B shown in white circles**, primers *A56R*-BVV-nDEL F and *A56R*-generic R were used.

3.2. A56R qPCR as a Tool to Study Pathogenesis and Viral Spread

3.2.1. Clinical Signs in Mice: Coinfected Versus Monoinfected

Mice that were inoculated with PV1-GII (1×10^6) and coinfected (PV1-GII + PV2-GI with 5×10^5 of each) showed the first clinical signs, such as facial edema, fur ruffling, hunching of the back, dyspnea and severe weight loss from 4 to 15 d.p.i. The coinfected mice presented a slightly longer clinical manifestation compared to the PV1-GII group, which was not significant with respect to percent weight loss (PV1-GII vs. PV1-GII + PV2-GI). In the PBS and PV2-GI groups, no clinical signs or weight loss were observed. The only significant differences were the weight variation of PV1-GII and the co-infected group compared to the PBS group (Figure 3).

3.2.2. Coinfected Mice Present Higher Frequency of *VACV* Detection in Lungs and Spleens than Monoinfected Groups

Using Reaction A, which amplifies all *VACV*, a positive qPCR was obtained in the lungs and spleen of 100 % of the mice inoculated with PV1- GII and PV1-GII + PV2-GI and in 60% of the PV2-GI infected mice lungs and 80% of the spleens. Reaction B of qPCR, which amplifies only the GII *A56R* gene (inoculated with PV1), showed a high positivity (100%) in lungs and only 60% of positivity in spleen (Figure 4A). This was confirmed by plaque phenotype assays (Figure 4B) of the lungs of four mice that were coinfected, which showed the presence of two viral populations (small plaques-PV2-GI and large plaques-PV1-GII).

3.2.3. Viral Load in Lung and Spleen

The analysis of the viral loads in the lungs of mice in the monoinfected group (PV1) revealed an average viral load of almost two logs higher than monoinfected group (PV2-GI) (Figure 5A). A significant difference was observed when the viral loads of both monoinfected groups (PV1-GII and PV2-GI) were compared with each sample in coinfection (PV1-GII with PV1-GII + PV2-GI and PV2-GI with PV1-GII + PV2-GI) (Figure 5A). Differences about one log for PV1-GII groups and almost three logs between PV2-GI mice groups were observed (Figure 5A).

The differences between the monoinfected groups (PV1-GII or PV2-GI) in the spleens were irrelevant, demonstrating a higher concentration of viral DNA of GI in this tissue (Figure 5B). The viral load in spleens of PV2-GI monoinfected was very similar to that of PV2-GI from the coinfected group (PV1 + PV2) (Figure 5B). When the viral load of PV1-GII monoinfected with PV1-GI from the coinfected group (PV1-GII + PV2-GI) was compared, it showed a decrease of about one log with statistical significance (Figure 5B).

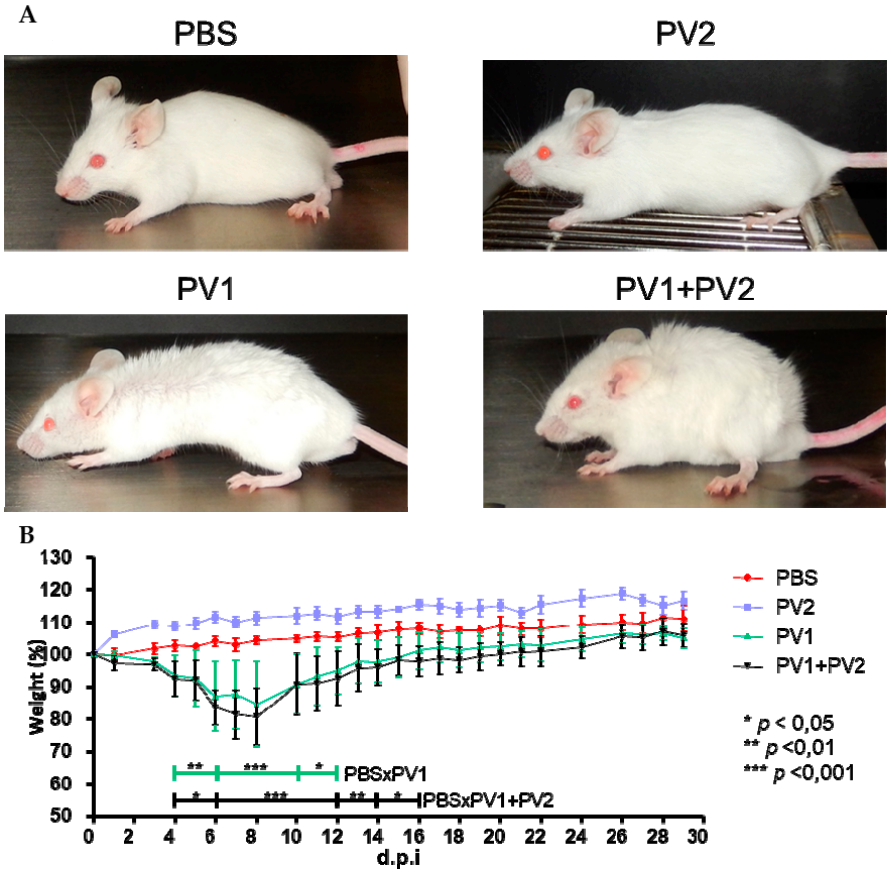


Figure 3. Clinical signs of Balb/c mice coinfected and monoinfected with PV1-GII and PV2-GI. Five-week-old male Balb/c mice were intranasally inoculated with 10 µL of viral suspension: PV1-GII 1×10^6 p.f.u.; PV2-GI 1×10^6 p.f.u.; PV1-GII + PV2-GI 5×10^5 p.f.u. of each sample; and a negative control group was inoculated with 10 µL of PBS. Clinical signs were observed on day 5 p.i and were recorded for 30 days post infection (d.p.i.). (**A**) In PBS and PV2-GI groups, no clinical signals were observed. Mice monoinfected with PV1-GII and coinfected (PV1-GII + PV2-GI) showed fur ruffling and hunching of the back. (**B**) The mice were daily weighed and relative mean weight was calculated. The error bars indicate standard deviations. PV1-GII: Pelotas virus 1 Group II; PV2-GI: Pelotas virus 2 Group I; PBS: phosphate-buffered saline. Asterisks indicate a statistically significant difference: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

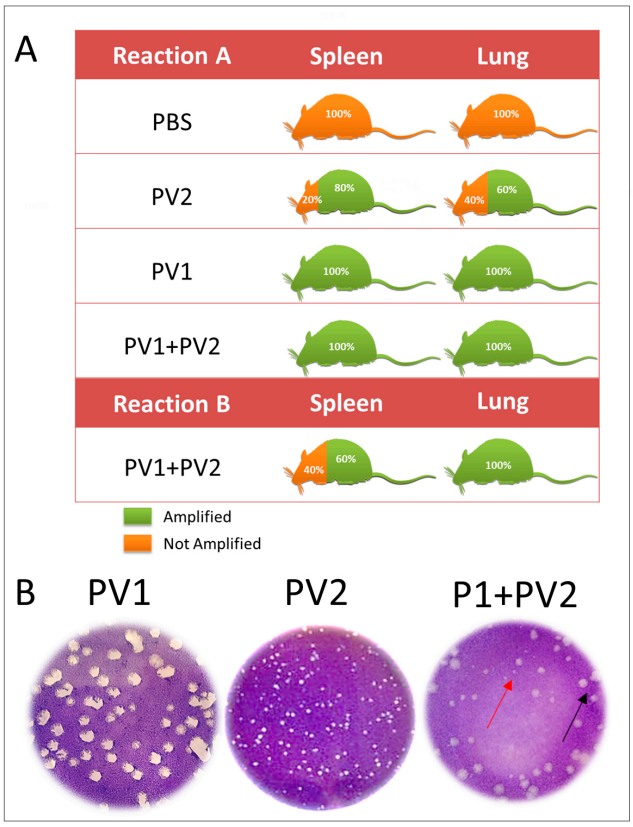


Figure 4. Detection of VAC-BR GI and GII in infected Balb/C mice: (**A**) qPCR detection of viral DNA from *VACV A56R* gene in mice spleen and lung samples. The results of Reaction A and Reaction B are represented. (**B**) Plaque phenotype assays of monoinfected (PV1-GII and PV2-GI) and coinfected mice showed two viral populations: small plaques-PV2 (red arrows) and large plaques-PV1 (black arrows). PV1-GII: Pelotas virus 1; PV2-GI: Pelotas virus 2; PBS: phosphate-buffered saline.

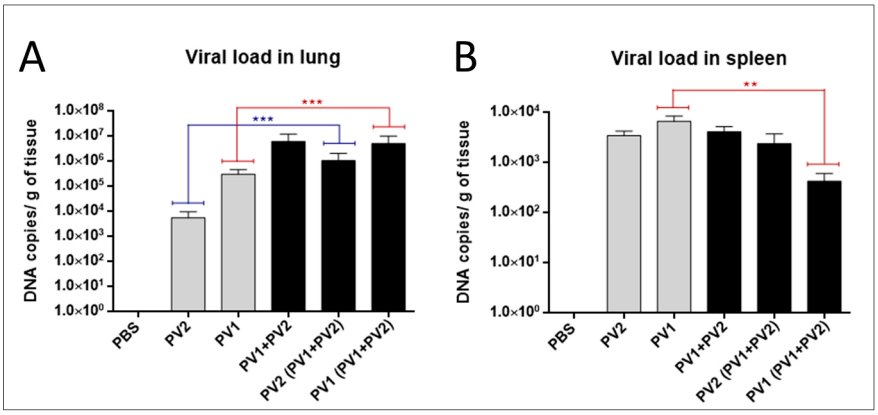


Figure 5. Viral load in mono or coinfected PV1-GII and PV2-GI mice. Viral loads in: lungs (**A**); and spleen (**B**) of Balb/c mice monoinfected (grey columns) and coinfected (black columns) with PV1-GII and PV2-GI were determined by qPCR. Comparisons between mono and coinfected groups are highlighted in red (PV1-GII groups) and blue (PV2-GI groups). The error bars indicate standard deviations. The statistical tests used were One-Way ANOVA and Tukey's multiple comparisons test Asterisks indicate a statistically significant difference: ** $p < 0.01$; *** $p < 0.001$. PV1: Pelotas virus 1; PV2: Pelotas virus 2; PBS: phosphate-buffered saline.

4. Discussion

Notifications of natural cases of *VACV* infection have increased [4,15,16], but this increase may be due to the development and improvement of study and detection techniques. qPCR is a rapid diagnostic tool, and is more sensitive than standard PCR; in addition, it allows acquiring molecular quantifications. Currently, multiplex real-time PCR has been described as a simple, reliable, and rapid method for the detection, identification and quantification of many kinds of viral coinfections [21–23]. In this way, qPCR assays have been used to detect and identify parapoxviruses and orthopoxviruses [24–27]. Here, we developed a new tool for the detection and quantification of *VACV*-BR isolates belonging to GI and GII during coinfections. This method was standardized under controlled conditions in animal model. Although the detection limit of 2 p.f.u./mg make its efficiency possible for many types of clinical samples more studies are needed to better clarify the total applications of method.

The qPCR assay to measure the viral load in coinfection was efficient and highly sensitive for different specimens (lung and spleen), showing a detection limit of 2 p.f.u./mg of tissue. The developed qPCR assay was useful as a detection system and in establishing the total viral load or relative viral load of *VACV*-GI and -GII in coinfections. The viral load could be measured because the efficiencies between two reactions are very close (96% and 94%).

Previous studies have shown that isolates of *VACV*-BR GII are virulent in murine model Balb/c, with weight loss and severe clinical signs, differently from GI, which is avirulent in this model [14]. Corroborating these data, the *VACV* isolates used in this work, in the monoinfected groups, PV1-GII and PV2-GI, followed the same pattern of virulence that has been previously described [15]. However, in this study, clinical signals in coinfected mice were evaluated and, unlike other studies conducted to date, this study evaluated animals until 30 d.p.i. when they had recovered from the infection [14–16]. Mice infected with the PV1-GII, a virulent representative of GII, and coinfected with PV2-GI, showed similar clinical signs, such as facial edema, fur ruffling, hunching of the back, dyspnea and severe weight loss. The weight variation data were significant at several points in the curve when we compared the control group to the groups infected with PV1-GII and co-infected (we use the two-way ANOVA test and Bonferroni method post-test).

The viral spread could be analyzed using this developed tool demonstrating a higher number of positive spleens samples (80%), compared to lung samples (60%) which is the primary site of inoculation in mice infected with PV2-GI. Furthermore, this analysis demonstrated the first description of spread of *VACV*-BR samples of GI. Viral loads data and the spread of *VACV*-BR GI isolates have not yet been demonstrated in other studies, and these results demonstrate the tendency of a higher tropism for lungs of PV1-GII and PV2-GI in coinfections.

In primary infection site (lungs), the coinfections with both viruses have higher viral loads, contrasting with monoinfected groups. On the other hand, in spleen the viral load of the virulent group (PV1-GII) is low in case of coinfection compared with the group monoinfected by same virus. These results may suggest an intriguing interaction among host and viruses which could help to understand why these two *VACV* groups are found in coinfections on nature. It is too early to answer the many questions raised by these results and more studies are needed to generalize these characteristics for all *VACV*-BR GI and GII coinfections, or as exclusive of PV1-GII and PV2-GI isolates.

Studies have shown that plaque size phenotypes are one of the biological characteristics that make possible the differentiation of the two groups of *VACV*-BR GI, with a small plaque phenotype, and GII with a large plaque phenotype [10,15,16,28]. Confirming biological characteristics of the plaque phenotype, the two viral populations (GI and GII) were detected in lungs of coinfected mice.

Using the developed tool of differential qPCR the first description of *VACV*-BR GI spread was possible. In previous studies, the GI virus could not be detected in mice lungs 5 d.p.i. [14]. Reports related to *VACV*-BR groups coinfections increases as well as the difficulty to detect both groups at the same time by methods of viral isolation and DNA sequencing. The main focus of this study was to develop an efficient tool to facilitate screening during outbreaks and consequently the study of coinfection, and also clarify its possible medical and veterinary importance.

Acknowledgments: We thank colleagues from Laboratório de Vírus (ICB-UFMG). This work was supported by research grants from Fundação de Amparo a Pesquisado Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). EGK and GST are researchers from CNPq.

Author Contributions: Conceived and designed the experiments: Rafael Calixto, Graziele Oliveira, Giliane de Souza Trindade, Danilo Bretas de Oliveira, and Erna Geessien Kroon. Performed the experiments: Rafael Calixto, Graziele Oliveira, Maurício Lima, and Ana Cláudia Andrade. Analyzed the data: Rafael Calixto, Danilo Bretas de Oliveira, and Erna Geessien Kroon. Contributed reagents/materials/analysis tools: Giliane de Souza Trindade and Erna Geessien Kroon. Wrote the paper: Rafael Calixto, Graziele Oliveira, Danilo Bretas de Oliveira, and Erna Geessien Kroon. All authors read and approved final manuscript.

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

References

1. Damon, I.K. Poxviruses. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, USA, 2014; Volume 2, pp. 2160–2184.
2. Moss, B. Poxviridae. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, USA, 2014; Volume 2, pp. 2129–2159.
3. Fenner, F.; Henderson, D.A.; Arita, I.; Jezek, A.; Ladnyi, I.D. *Smallpox and Its Eradication*; World Health Organization Press: Geneva, Switzerland, 1988.
4. Trindade, G.S.; Lobato, Z.I.; Drumond, B.P.; Leite, J.A.; Trigueiro, R.C.; Guedes, M.I.; da Fonseca, F.G.; dos Santos, J.R.; Bonjardim, C.A.; Ferreira, P.C.; et al. Short report: Isolation of two vaccinia virus strains from a single bovine vaccinia outbreak in rural area from Brazil: Implications on the emergence of zoonotic orthopoxviruses. *Am. J. Trop. Med. Hyg.* **2006**, *75*, 486–490. [PubMed]
5. Singh, R.K.; Hosamani, M.; Balamurugan, V.; Bhanuprakash, V.; Rasool, T.J.; Yadav, M.P. Buffalopox: An emerging and re-emerging zoonosis. *Anim. Health Res. Rev.* **2007**, *8*, 105–114. [CrossRef] [PubMed]
6. Kroon, E.G.; Mota, B.E.; Abrahão, J.S.; da Fonseca, F.G.; de Souza Trindade, G. Zoonotic Brazilian Vaccinia virus: From field to therapy. *Antivir. Res.* **2011**, *92*, 150–163. [CrossRef] [PubMed]
7. Oliveira, D.B.; Assis, F.L.; Ferreira, P.C.; Bonjardim, C.A.; de Souza Trindade, G.; Kroon, E.G.; Abrahão, J.S. Group 1 Vaccinia virus zoonotic outbreak in Maranhao State, Brazil. *Am. J. Trop. Med. Hyg.* **2013**, *89*, 1142–1145. [CrossRef] [PubMed]
8. Franco-Luiz, A.P.; Fagundes-Pereira, A.; Costa, G.B.; Alves, P.A.; Oliveira, D.B.; Bonjardim, C.A.; Ferreira, P.C.; de Souza Trindade, G.; Panei, C.J.; Galosi, C.M.; et al. Spread of vaccinia virus to cattle herds, Argentina, 2011. *Emerg. Infect. Dis.* **2014**, *20*, 1576–1578. [CrossRef] [PubMed]
9. Damaso, C.R.; Esposito, J.J.; Condit, R.C.; Moussatché, N. An emergent poxvirus from humans and cattle in Rio de Janeiro State: Cantagalo virus may derive from Brazilian smallpox vaccine. *Virology* **2000**, *277*, 439–449. [CrossRef] [PubMed]
10. Trindade, G.S.; Emerson, G.L.; Carroll, D.S.; Kroon, E.G.; Damon, I.K. Brazilian vaccinia viruses and their origins. *Emerg. Infect. Dis.* **2007**, *13*, 965–972. [CrossRef] [PubMed]
11. Megid, J.; Appolinário, C.M.; Langoni, H.; Pituco, E.M.; Okuda, L.H. Vaccinia virus in humans and cattle in southwest region of Sao Paulo state, Brazil. *Am. J. Trop. Med. Hyg.* **2008**, *79*, 647–651. [PubMed]
12. Assis, F.L.; Almeida, G.M.; Oliveira, D.B.; Franco-Luiz, A.P.; Campos, R.K.; Guedes, M.I.; Fonseca, F.G.; Trindade, G.S.; Drumond, B.P.; Kroon, E.G.; et al. Characterization of a new Vaccinia virus isolate reveals the C23L gene as a putative genetic marker for autochthonous Group 1 Brazilian Vaccinia virus. *PLoS ONE* **2012**, *7*, e50413. [CrossRef] [PubMed]
13. Drumond, B.P.; Leite, J.A.; da Fonseca, F.G.; Bonjardim, C.A.; Ferreira, P.C.; Kroon, E.G. Brazilian Vaccinia virus strains are genetically divergent and differ from the Lister vaccine strain. *Microbes Infect.* **2008**, *10*, 185–197. [CrossRef] [PubMed]
14. Ferreira, J.M.; Drumond, B.P.; Guedes, M.I.; Pascoal-Xavier, M.A.; Almeida-Leite, C.M.; Arantes, R.M.; Mota, B.E.; Abrahão, J.S.; Alves, P.A.; Oliveira, F.M.; et al. Virulence in murine model shows the existence of two distinct populations of Brazilian Vaccinia virus strains. *PLoS ONE* **2008**, *3*, e3043. [CrossRef] [PubMed]

15. Campos, R.K.; Brum, M.C.; Nogueira, C.E.; Drumond, B.P.; Alves, P.A.; Siqueira-Lima, L.; Assis, F.L.; Trindade, G.S.; Bonjardim, C.A.; Ferreira, P.C.; et al. Assessing the variability of Brazilian Vaccinia virus isolates from a horse exanthematic lesion: Coinfection with distinct viruses. *Arch. Virol.* **2011**, *156*, 275–283. [CrossRef] [PubMed]
16. Oliveira, G.; Assis, F.; Almeida, G.; Albarnaz, J.; Lima, M.; Andrade, A.C.; Calixto, R.; Oliveira, C.; Diomedes Neto, J.; Trindade, G.; et al. From lesions to viral clones: Biological and molecular diversity amongst autochthonous Brazilian vaccinia virus. *Viruses* **2015**, *7*, 1218–1237. [CrossRef] [PubMed]
17. Leite, J.A.; Drumond, B.P.; de Souza Trindade, G.; Bonjardim, C.A.; Ferreira, P.C.; Kroon, E.G. Brazilian Vaccinia virus strains show genetic polymorphism at the ati gene. *Virus Genes* **2007**, *35*, 531–539. [CrossRef] [PubMed]
18. Assis, F.L.; Borges, I.A.; Ferreira, P.C.; Bonjardim, C.A.; de Souza Trindade, G.; Lobato, Z.I.; Guedes, M.I.; Mesquita, V.; Kroon, E.G.; Abrahão, J.S. Group 2 vaccinia virus, Brazil. *Emerg. Infect. Dis.* **2012**, *18*, 2035–2038. [CrossRef] [PubMed]
19. Brum, M.C.S.; Anjos, B.L.; Nogueira, C.E.W.; Amaral, L.A.; Weiblen, R.; Flores, E.F. An outbreak of orthopoxvirus-associated disease in horses in southern Brazil. *J. Vet. Diagn. Investig.* **2010**, *22*, 143–147.
20. Trindade, G.S.; Li, Y.; Olson, V.A.; Emerson, G.; Regnery, R.L.; da Fonseca, F.G.; Kroon, E.G.; Damon, I.K. Real-time PCR assay to identify variants of Vaccinia virus: Implications for the diagnosis of bovine vaccinia in Brazil. *J. Virol. Methods* **2008**, *152*, 63–71. [CrossRef] [PubMed]
21. Zheng, L.L.; Wang, Y.B.; Li, M.F.; Chen, H.Y.; Guo, X.P.; Geng, J.W.; Wang, Z.Y.; Wei, Z.Y.; Cui, B.A. Simultaneous detection of porcine parvovirus and porcine circovirus type 2 by duplex real-time PCR and amplicon melting curve analysis using SYBR Green. *J. Virol. Methods* **2013**, *187*, 15–19. [CrossRef] [PubMed]
22. Venkatesan, G.; Balamurugan, V.; Bhanuprakash, V. TaqMan based real-time duplex PCR for simultaneous detection and quantitation of capripox and orf virus genomes in clinical samples. *J. Virol. Methods* **2014**, *201*, 44–50. [CrossRef] [PubMed]
23. Balboni, A.; Dondi, F.; Prosperi, S.; Battilani, M. Development of a SYBR Green real-time PCR assay with melting curve analysis for simultaneous detection and differentiation of canine adenovirus type 1 and type 2. *J. Virol. Methods* **2015**, *222*, 34–40. [CrossRef] [PubMed]
24. Carletti, F.; di Caro, A.; Calcaterra, S.; Grolla, A.; Czub, M.; Ippolito, G.; Capobianchi, M.R.; Horejsh, D. Rapid, differential diagnosis of orthopox- and herpesviruses based upon real-time PCR product melting temperature and restriction enzyme analysis of amplicons. *J. Virol. Methods* **2005**, *129*, 97–100. [CrossRef] [PubMed]
25. Fedele, C.G.; Negredo, A.; Molero, F.; Sánchez-Seco, M.P.; Tenorio, A. Use of internally controlled real-time genome amplification for detection of variola virus and other orthopoxviruses infecting humans. *J. Clin. Microbiol.* **2006**, *44*, 4464–4470. [CrossRef] [PubMed]
26. Nitsche, A.; Büttner, M.; Wilhelm, S.; Pauli, G.; Meyer, H. Real-time PCR detection of parapoxvirus DNA. *Clin. Chem.* **2006**, *52*, 316–319. [CrossRef] [PubMed]
27. Li, Y.; Olson, V.A.; Laue, T.; Laker, M.T.; Damon, I.K. Detection of monkeypox virus with real-time PCR assays. *J. Clin. Virol.* **2006**, *36*, 194–203. [CrossRef] [PubMed]
28. Abrahão, J.S.; Guedes, M.I.; Trindade, G.S.; Fonseca, F.G.; Campos, R.K.; Mota, B.F.; Lobato, Z.I.; Silva-Fernandes, A.T.; Rodrigues, G.O.; Lima, L.S.; et al. One more piece in the VACV ecological puzzle: Could peridomestic rodents be the link between wildlife and bovine vaccinia outbreaks in Brazil? *PLoS ONE* **2009**, *4*, e7428. [CrossRef] [PubMed]

frontiers in Microbiology

ORIGINAL RESEARCH
published: 30 August 2017
doi: 10.3389/fmicb.2017.01673

An Anthropocentric View of the Virosphere-Host Relationship

*Rodrigo A. L. Rodrigues, Ana C. dos S. P. Andrade, Paulo V. de M. Boratto, Giliane de S. Trindade, Erna G. Kroon and Jônatas S. Abrahão**

Laboratório de Vírus, Department of Microbiology, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

For over a century, viruses have been known as the most abundant and diverse group of organisms on Earth, forming a virosphere. Based on extensive meta-analyses, we present, for the first time, a wide and complete overview of virus–host network, covering all known viral species. Our data indicate that most of known viral species, regardless of their genomic category, have an intriguingly narrow host range, infecting only 1 or 2 host species. Our data also show that the known virosphere has expanded based on viruses of human interest, related to economical, medical or biotechnological activities. In addition, we provide an overview of the distribution of viruses on different environments on Earth, based on meta-analyses of available metaviromic data, showing the contrasting ubiquity of head-tailed phages against the specificity of some viral groups in certain environments. Finally, we uncovered all human viral species, exploring their diversity and the most affected organic systems. The virus–host network presented here shows an anthropocentric view of the virology. It is therefore clear that a huge effort and change in perspective is necessary to see more than the tip of the iceberg when it comes to virology.

Keywords: virosphere, anthropocentric, virus–host relationship, network, metavirome

OPEN ACCESS

Edited by:
William Michael McShan, University of Oklahoma Health Sciences Center, United States

Reviewed by:
Juliana Felipetto Cargnelutti, Universidade Federal de Santa Maria, Brazil
Jessica Labonté, Texas A&M University at Galveston, United States

****Correspondence:***
Jônatas S. Abrahão
jonatas.abrahao@gmail.com

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *11 July 2017*
Accepted: *17 August 2017*
Published: *30 August 2017*

Citation:
Rodrigues RAL, Andrade ACSP, Boratto PVM, Trindade GS, Kroon EG and Abrahão JS (2017) An Anthropocentric View of the Virosphere-Host Relationship. Front. Microbiol. 8:1673. doi: 10.3389/fmicb.2017.01673

INTRODUCTION

The virology, as a science field, started at the end of the XIX century with the studies of Adolf Mayer, Dmitry Ivanofsky, and Martinus Beijerinck about tobacco mosaic disease. The investigators noticed that they were dealing with an agent completely unknown to the academic community, which retained its infectious nature even after passing through Chamberland filters (at that time, the most efficient method to retain bacteria). Furthermore, even after being diluted by filtration in a porous membrane, the agent recovered its infectiveness after replication within living tissues of healthy plants. The new pathogen was named "*contagium vivum fluidum*," and only after the advent of *in vitro* plaque assays and electron microscopy it was fully recognized as a virus (Enquist and Racaniello, 2013). Lwoff (1957) published a seminal work in which he established, for the first time, a set of characteristics for an organism to be considered a virus; among them were being an intracellular parasite and completely relying on the biosynthetic machinery of its host, thus being considered a non-living organism. With the advancement of virology, the International Committee on Taxonomy of Viruses (ICTV) was created in the 1960s (originally the International Committee

on Nomenclature of Viruses) with the objective of cataloging and organizing the viruses that were being described in the years to come; it established the first rules for viral taxonomy. A few years later, David Baltimore proposed a strategy to organize the viruses according to the properties of their genetic material, with six groups being defined at that time: I (dsDNA), II (ssDNA), III (dsRNA), IV [ssRNA(+)], V [(ssRNA(−)], and VI (ssRNA-RT) (Baltimore, 1971). In the following years, two additional groups were considered, composing the groups VII (dsDNA-RT) and VIII (viroids). This organization strategy is currently well accepted among virologists.

In the years to come, several viruses were described, being isolated in every corner of the planet from hosts belonging to the three domains of life, i.e., Eukarya, Bacteria, and Archaea. In this context, the virus species concept was created by the ICTV, which is the lowest taxon (group) in a branching hierarchy of viral taxa, defined as a polythetic class of viruses that constitute a replicate lineage and occupy a particular ecological niche (i.e., possess similar biological features) (International Committee on Taxonomy of Viruses - Taxonomy, 2017). These viruses continuously reaffirmed the established criteria raised in the 1950s to recognize an organism as a virus. Only during the last few years this paradigm was broken with the discovery of giant viruses (La Scola et al., 2003; Boyer et al., 2009; Philippe et al., 2013; Legendre et al., 2014). These viruses put the well-established concepts to the test, restoring debates about their complete dependency on their hosts and whether they should be considered living organisms, therefore deserving a place in the metaphorical tree of life (Raoult and Forterre, 2008; Forterre, 2010). Besides, advancements in the field of genomics during the last few years, especially metagenomics (or even metaviromics), have allowed the identification of countless viral sequences in several regions of the globe, supporting previous electron microscopy data which suggested the viral ubiquity and an astronomical number of viruses on Earth, thus forming a virosphere (Suttle, 2005; Kristensen et al., 2010).

Although the identification of new viruses and studies of their interaction with hosts have considerably advanced, we still do not know how this interactive network is truly connected. Moreover, many metaviromic studies have been developed allowing the identification of different viral sequences around the world, but we do not have a clear vision of how the viral diversity is distributed on the planet, or how much we have searched for new viruses. Therefore, a new look into what is currently available and the use of new strategies to explore these data could bring new insights and allow the advancement of the virology field. Through extensive meta-analysis of currently available data, we demonstrate here that the known viruses have a very narrow host range, resulting in a spatially connected network. We found a highly anthropocentric view of the virosphere and demonstrated the existence of some specific viral groups in certain environments on the Earth, leading us to reflect about how far we have progressed in the study of viruses. Finally, we analyzed the diversity of human-associated viruses and the tropism of these viruses. The results presented here show a highly biased virology, confirming that we know only the tip of the iceberg and a lot of work remains to be done so we can have a clearer view of the diversity and ecology of the virosphere.

MATERIALS AND METHODS

Dataset Preparation and Selection Criteria

Virosphere and Hosts

To analyze the host range of the known viruses, only those officially recognized by the International Committee on Taxonomy of Viruses (ICTV) were included in the analysis. The definition of the best dataset to perform this analysis comprises a challenging task. In this context, ICTV proved to be the best option for gathering the largest and most updated dataset of recognized virus species, grouping and reflecting the diversity and circulation of viruses in nature. A list containing all of the virus species was downloaded from ICTV website[1]. A list released on May 26th, 2016 was used. Therefore, new viruses classified by means of metagenomic data, following the new criteria recently approved by the Executive Committee of ICTV (Simmonds et al., 2017), as wells as the reclassification of the family *Bunyaviridae,* were not considered in this analysis. We considered hosts those organisms in which we found consistent and recurrent evidences of the detection of a virus in a given species by means of isolation, serology, and molecular detection. This detection was associated in most cases with clinical manifestation and, in a few cases, in a non-disease context. Organisms used as study models were not considered here. Hosts were associated with each virus at the lowest taxonomic level possible using the Virus–Host Database (Mihara et al., 2016), VIDE database[2], and full research articles related to a given virus. In the latter, only one reference was used to determine the host species, even though more than one study (whenever available) was analyzed to corroborate the reference used. During our research and analyses, we considered (whenever the data were available) different viruses within a virus species and their host-range. Only the viruses in which it was possible to determine the hosts at species or genus taxonomic level were considered for the construction of the network. A total of 4497 nodes were included in the network dataset, classified as virus, animalia, plantae, fungi, protist, bacteria, and archaea, along with 4814 edges directly connecting the nodes, all with weight $(w) = [1]$.

Viral Diversity

To analyze the known viral diversity on the planet, we considered viral groups (families recognized by the ICTV or groups currently unassigned to a proper taxa) identified in diverse metavirome studies performed in the following environments: marine [10], freshwater [7], soil [6], hypersaline [5], thermal springs [4], sewage [4], and polar water [3], in a total of 39 works. The studies were accessed at National Center for Biotechnology Information

[1]https://talk.ictvonline.org/files/master-species-lists/

[2]http://sdb.im.ac.cn/vide/sppindex.htm

(NCBI)[3] using the name of the environments added by virome or metavirome as keywords in the search field. All of the viral groups identified were included in the network analysis, where they were associated with the environments in which they were detected. A total of 103 nodes were included in the network graph, classified according to the analyzed environments and viral order recognized by the ICTV [*Ligamenvirales, Tymovirales, Herpesvirales, Caudovirales, Picornavirales, Mononegavirales, Nidovirales*, and those not classified in order (Unassigned)], and 260 edges indirectly connecting the nodes, with $w = [1]$. To better visualize the viral groups shared between different environments, we created a circular layout image using Circos package (Krzywinski et al., 2009). In addition to the detected viral groups, we computed the type of technology used for nucleic acid sequencing, the type of material analyzed (DNA or RNA), and whether a 200 nm filter was used for sample preparation.

Human Viruses and Viral Tropism

The viruses that affect humans were defined after the association of the hosts of each virus species recognized by the ICTV, as described above. The viruses were associated with the following organic systems, according to the clinical manifestation reported in cases of infection: digestive, integumentary, respiratory, nervous, muscular, skeletal, cardiovascular, urinary, reproductive, lymphatic, immune, endocrine, or none of them, in cases of non-pathogenic viruses, based on clinical manifestation and/or tropism for a particular body tissue. Clinical manifestation and the tropism for each system were defined according to full research articles found at NCBI and using the arboviruses catalog of the Center for Disease Control and Prevention[4]. The viruses were associated with different systems in a bipartite network composed of 333 nodes classified according to the organic systems and viruses, and 497 edges indirectly connecting the nodes, with $w = [1]$. In parallel, we built a unipartite network graph wherein the systems were interconnected according to the viruses that affect different systems simultaneously, in a total of 12 nodes and 42 edges indirectly connecting the nodes, with $w = [1,25]$.

Construction of Networks

The networks presented in this work were built using the program Gephi version 0.9.1 (Bastian et al., 2009). All components of the each graph were listed in a comma-separated values (.csv) spreadsheet, which was imported to the software. Another .csv spreadsheet containing the connections between the components was also imported to generate the raw graph. In all networks, the node diameter is directly proportional to the edge degree. The thickness of the edges is directly proportional to the number of times that a node is connected to another, wherein different weights were assigned to the edges. The layout was generated using algorithms based on force of attraction and repulsion of the nodes (Fruchterman-Reingold followed by ForceAtlas 2), followed by local rearrangement of the nodes for a better visualization of the connections between nodes, without perturbing the general layout of the networks.

[3]https://www.ncbi.nlm.nih.gov/pubmed/

[4]https://wwwn.cdc.gov/arbocat/

RESULTS AND DISCUSSION

The Known Viruses Have a Very Narrow Host Range

The ICTV is the organization responsible for cataloging and classifying viruses into virus species that have been described over time. Historically, this organization has taken into consideration several criteria for a new isolate to be considered a new species, such as the genetic material and the hosts in which it was isolated, as well as any clinical manifestations it may possibly cause (Simmonds et al., 2017). Viral taxonomy covers the levels of order, family (and subfamily in some cases), genus and species, wherein the vast majority of virus species remain outside of a virus order. All of this information is constantly updated by the ICTV, which periodically publishes the Master Species List (MSL). In this work, we evaluated the host range of all known viruses with a virus species officially recognized and published by the ICTV on May 26th, 2016 (MSL#30) [**Supplementary Table S1**]. An extensive search using public databases and indexed publications was performed to define the natural hosts of all of the viruses present in the list (see Materials and Methods). The majority of the viruses present in the MSL#30 (a total of 3704 virus species, henceforward named the known virosphere) comprises group I (dsDNA) and IV [ssRNA(+)] according to Baltimore's classification [35 and 28%, respectively, followed by group II (ssDNA – 17%)], with the remaining groups representing 20% of the known virosphere (**Figure 1A**). It was possible to associate hosts at the species or genus level to 3414 viruses (92.2%), at the family level or higher to 265 viruses (7.15%), and it was not possible to associate any host for only 25 viruses (0.65%), either because the natural hosts for the viruses are not yet known, or due to a complete lack of information in the literature about their host range (**Figure 1B**). For all viral groups, according to Baltimore's classification, the host range is very restricted, with more than 50% of known viruses infecting only one or two host species, reaching up to 75% in some groups, such as those viruses with genomes composed of dsDNA, ssDNA, ssRNA-RT, and viroids (**Figure 1C**). Only the ssRNA(−) viruses seems to possess a slightly broader host range, wherein 42% of the viruses are able to infect more than four host species. Considering the entire known virosphere, 73.3% are associated with only one or two host species; 3.5% with three or four species; 22.5% with more than four species; and only 0.7% have a natural host range which has not been defined (**Figure 1C**). These analyses reveal that, until now, based on the available information we have, viruses have a very narrow host range. This disturbing data must be interpreted carefully. It is likely that several unknown viruses have a broader host-range, which will drastically change the view presented here; however, we might be far from acquire this kind of knowledge since these relationships are likely out of scope of human investigation. Therefore, in light of the research performed so far, we are facing such suspicious data.

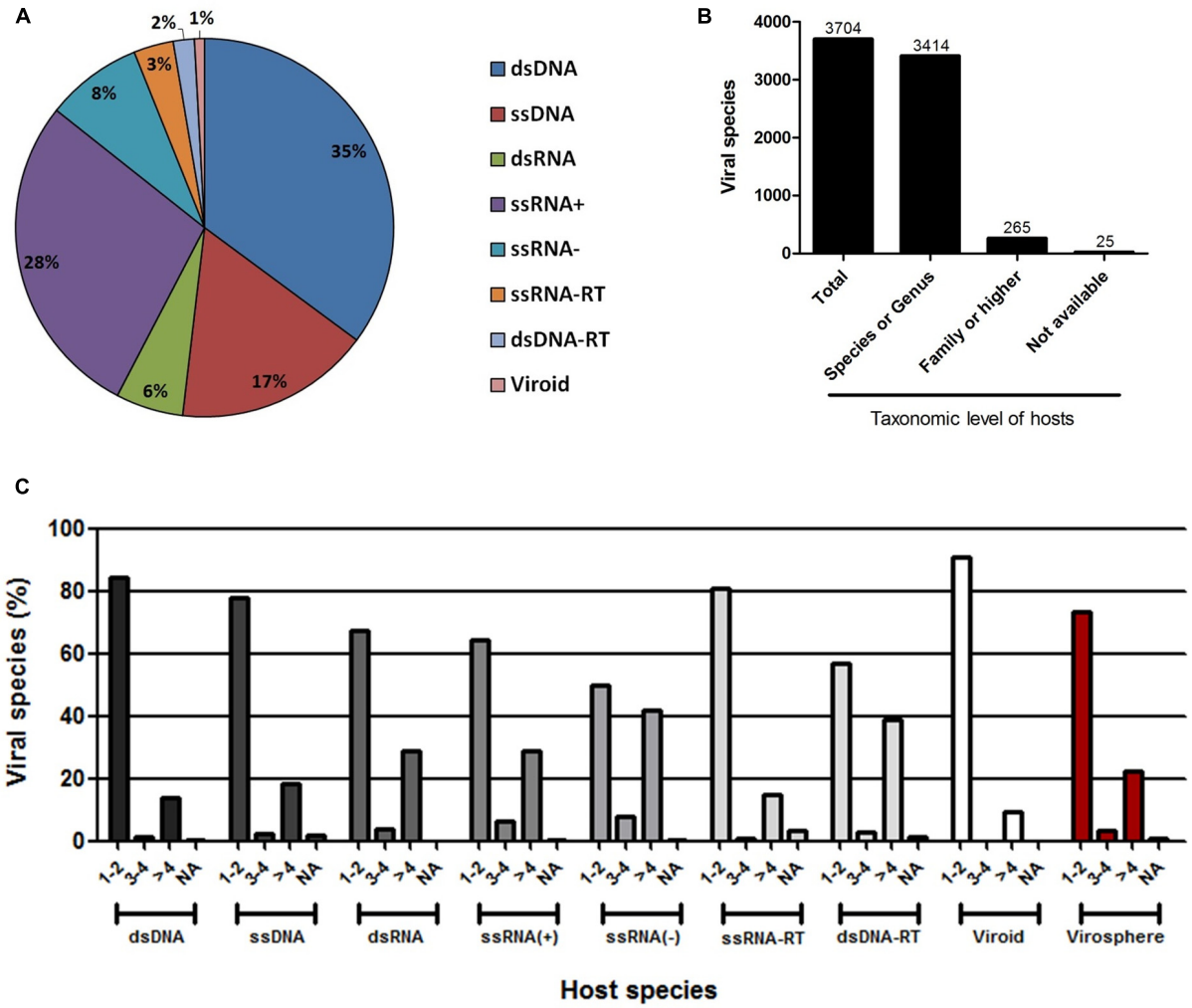


FIGURE 1 | Host range of the known virosphere. **(A)** Pie chart showing the distribution of the viruses recognized by the International Committee on Taxonomy of Viruses (ICTV) according to Baltimore's classification. **(B)** Taxonomic level of the hosts associated to the known viruses. More than 90% of the viruses were associated to hosts at species or genus taxonomic level, which were used in following analysis. **(C)** Amount of host species for viruses according to Baltimore's classification, showing a very narrow host range of the viruses. NA, not available.

An Anthropocentric View of the Known Virosphere

To better represent the interaction between the viruses and the hosts so that we can have a clear vision of how interconnected these organisms are, we built a bipartite network graph composed of 4497 nodes, with 3414 viruses (only viruses associated with hosts at species or genus taxonomic level were included in this analysis) and 1083 hosts (at genus level), all connected by 4814 edges with the same weight (w) = [1]. The hosts were classified according to the major realms and domains of life: Animalia, Plantae, Protist, Fungi, Bacteria, and Archaea (Woese, 2002). We observed a spatially connected network, wherein only a few hosts were associated to a huge amount of viruses, while the majority of the hosts are associated with a few viruses, a reflex of the very narrow host range of the known virosphere (**Figure 2**). Furthermore, the analysis of the network revealed a highly anthropocentric virosphere, in which most viruses are associated with humans or hosts that are directly related to humans by economic, medicinal or biotechnological interests. The vast majority of known viruses are associated with plants (483 genera) or animals (467 genera). These groups are more interconnected than others, even though more than 70% of these hosts possess only one or two associated viruses (**Supplementary Figure S1**). It is noteworthy that some viruses can cross broad host categories, infecting both plants and animals. These viruses are plant pathogens transmitted by arthropod vectors, in which are able to fully replicate and reach the plant host (Dietzgen et al., 2016). Bacteria-infecting viruses (known as bacteriophages or phages) are mainly distributed among the families *Myoviridae*, *Podoviridae*, and *Siphoviridae* (order *Caudovirales*), and are associated with 62 known host genera. This group is spatially connected, reflecting the narrow host range of phages. However, different to animals and plants, almost 40% of known bacteria

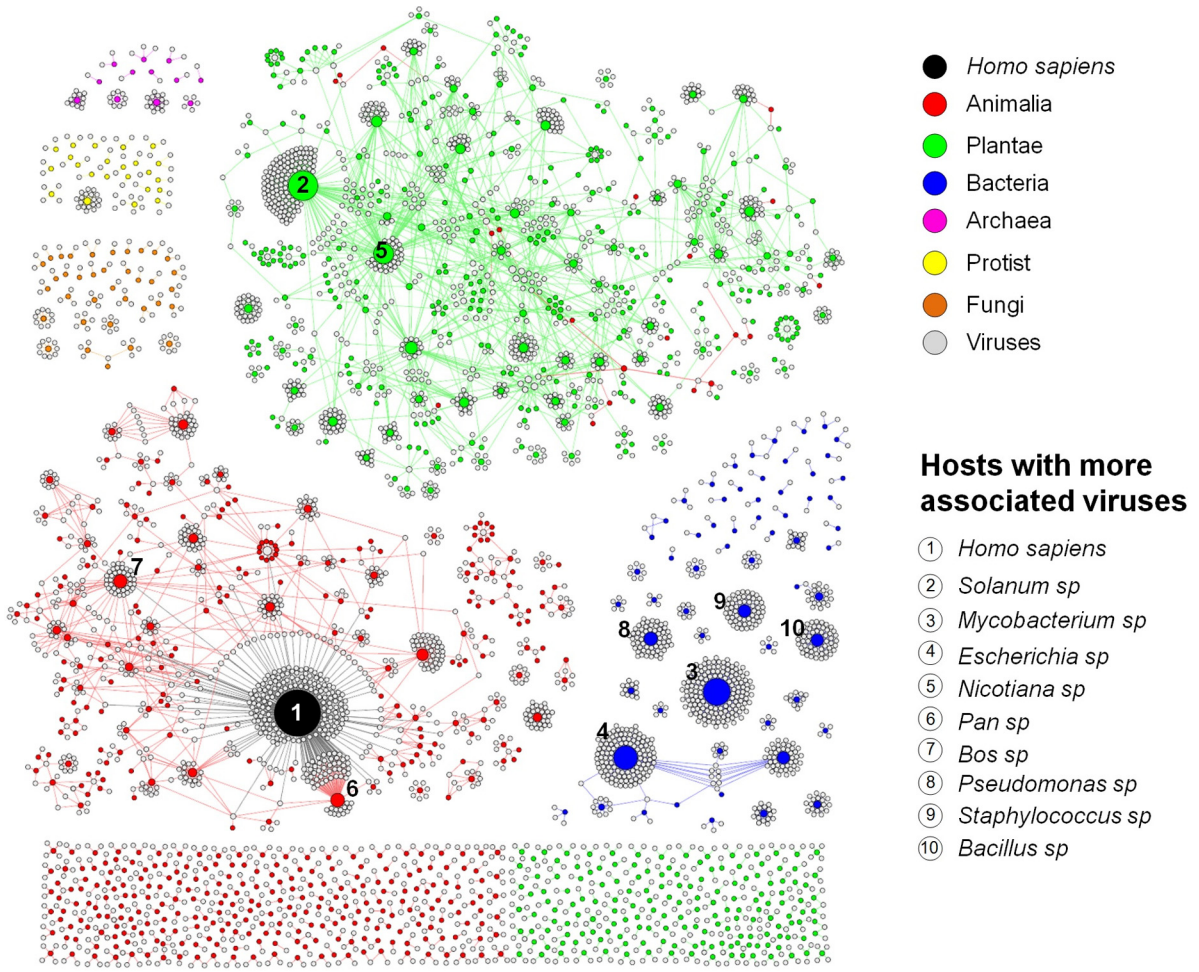


FIGURE 2 | Virus–host interaction network. Bipartite network graph showing a spatially connected network among viruses and hosts, a reflection of our limited knowledge about the viruses and their hosts. Each node represents a virus (gray), or a host genus, classified according to the taxonomic group (colored nodes). The nodes' diameter is proportional to the edge degree. The layout was generated using a force based algorithm followed by manual rearrangement to a better visualization of the connections. A total of 3414 viruses and 1083 hosts (genus level) are represented. The 10 hosts with more associated viruses are depicted.

are infected by more than four viruses. Some bacteria comprised hubs in the network, such as *Mycobacterium* and *Escherichia*, with several associated viruses. Since they are intensively studied due to their medicinal and biotechnological relevance (Korb et al., 2016; Vila et al., 2016), it was expected that a large number of viruses would be identified as parasites of these groups. In fact, a large majority of phage sequences available in GenBank was isolated from a few groups of bacteria associated to human diseases or food processing (Holmfeldt et al., 2013). The knowledge about viruses affecting fungi, protists and archaea is scarce, probably due to the lack of investigation of these groups of viruses and their hosts. These viruses were associated with 36 genera of fungi, 23 protists, and only 12 genera of archaea, reflecting how poorly these microorganisms are studied under the lens of virology.

Among the host genera of each group that possess more associated viruses, many are composed of domesticated species such as *Bos* sp., *Sus* sp., and *Gallus* sp. (Animalia; e.g., cattle, swine, and chickens, respectively); *Solanum* sp., *Nicotiana* sp., *Phaseolus* sp., *Capsicum* sp., and *Cucumis* sp. (Plantae; e.g., potato, tobacco, common bean, peppers, and cucumber, respectively); *Chlorella* sp. (Protist); and *Saccharomyces* sp. (Fungi) (**Supplementary Figure S2**). Many species of these groups are employed in farming, such as cattle, pigs and poultry, as well as many grains and legumes consumed worldwide, handling billions of dollars annually (Thornton, 2010; Reganold and Wachter, 2016). In addition, some species of green algae (*Chlorella* sp., *Chlorophyta* phylum) are used as dietary supplementation as sources of vitamins and macronutrients and its efficacy against some human diseases are under constant investigation (Ebrahimi-Mameghani et al., 2016; Panahi et al., 2016). Yeasts of the *Saccharomyces* genus, especially *S. cerevisiae*, are considered domesticated fungi, being used worldwide in the production of alcoholic beverages, also making them economically important (Sicard and Legras, 2011; Gallone et al., 2016). Given the economic relevance of these organisms,

constant efforts are made to reveal parasites that might be considered a threat to them, thus enabling possible strategies of control and prevention to be established. Therefore, it was expected that these groups of hosts had more known viruses.

Other hosts are known due to their medicinal relevance for humans or animals and commercially explored plants, such as *Acanthamoeba* sp. and *Trichomonas* sp. (Protist), both related to severe infections in humans (Siddiqui and Khan, 2012; Menezes et al., 2016); *Heterobasidion* sp., *Cryphonectria* sp., *Rosellinia* sp., and *Ophiostoma* sp. (Fungi), groups of fungi related to diverse plant infections, both domesticated and from native forests, causing severe diseases such as annosum root and chestnut blight (Hillman and Suzuki, 2004; Ďurkovič et al., 2013; Kondo et al., 2013; Vainio and Hantula, 2015); and *Mycobacterium* sp., *Escherichia* sp., *Pseudomonas* sp., *Staphylococcus* sp., and *Bacillus* sp. (Bacteria), all groups of prokaryotes related to life-threatening diseases, such as tuberculosis (Korb et al., 2016), gastrointestinal, respiratory and urinary infections (Langan et al., 2015; Vila et al., 2016), and also used as biological weapons (Goel, 2015). Therefore, it is expected that these species are the target of intense investigation, and the majority of known phages are associated with these bacteria. Finally, some hosts are important in the biotechnology field or used as laboratory study models for molecular biology, such as *Ectocarpus* sp. (Protist) (Lipinska et al., 2016); *Sulfolobus* sp., and *Thermus* sp. (Archaea) (Cava et al., 2009; Zhang et al., 2013) (**Supplementary Figure S2**). Altogether, the data presented here show that in all group of hosts, both eukaryotic and prokaryotic, most of the known viruses are related to hosts that are important for humans in certain aspects. In this way, the virus–host network shows a highly anthropocentric view of the virology performed so far. This biased virology is probably the very reason for our view of a narrow host-range of the known viruses.

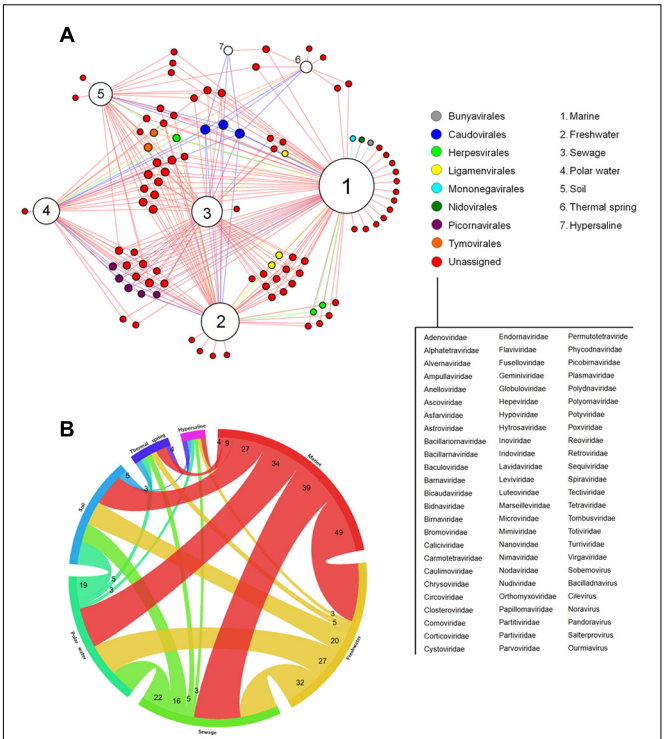


FIGURE 3 | Viral diversity on Earth. **(A)** Network graph showing the viral groups detected in different environments by metaviromic analysis. Each node represents an environment (white) or viral groups (families or known viral genus – colored nodes) classified according to the orders formally recognized by the ICTV. The viruses not currently assigned in any order are listed. The node diameter is proportional to the edge degree. The layout was generated using a force based algorithm followed by manual rearrangement for a better visualization of the connections. A total of 96 viral groups are represented. **(B)** Relationship between the different environments based on the amount of shared viral groups.

Viral Diversity on Earth

Since the discovery of the tobacco mosaic virus at the end of XIX century, many other viruses have been described and biologically characterized in many regions of the planet, thus contributing to the concept of viral ubiquity. With advances in electron microscopy techniques, many studies have been conducted in order to define the abundance and diversity of viruses, coming to an astronomic number, in the order of 10^{31} viral particles on the Earth (Suttle, 2005). However, only with the advent of massive parallel sequencing of nucleic acids and the development of a new research field – metagenomics – it was possible to create a better view of the viral diversity on the planet, reaffirming the viral ubiquity concept (Kristensen et al., 2010).

By analyzing different available metagenomic works, more specifically metaviromic works (analysis of viral nuclei acid sequences in different environments), we built a bipartite network graph connecting the viral groups found within seven distinct environments around the planet: marine, freshwater, polar water, thermal springs, hypersalines, and sewage (**Figure 3A**). A total of 39 works were analyzed (for choice criteria, see Materials and Methods). A total of 96 viral groups (genus or family) were detected in those studies. Different amount of viral groups are shared among the environments, wherein marine shared up to 49 viral groups with other environments, reinforcing the ubiquity of viruses on the planet (**Figure 3B**). Among the viral groups identified, only representatives of the families *Myoviridae*, *Podoviridae*, and *Siphoviridae* (phages belonging to the order *Caudovirales*) were found in all of the searched environments. After the initial studies of metagenomics in marine environments, in which they searched basically for bacteriophages, the hypothesis "Everything is everywhere but environment selects" was applied to these viruses, stating the ubiquity of the phages, even though some groups were specifically found in certain environments (O'Malley, 2008; Thurber, 2009). Our meta-analysis corroborates this hypothesis and goes further, showing that head-tailed phages are found in every location investigated, not only in marine samples. In contrast, the majority of viral groups were found only in two or three environments, and surprisingly, some groups were also restricted to only one environment (**Figure 3A**). The viral diversity is higher in marine environments, wherein 15 groups were exclusive to it. The great diversity of viruses in the oceans is a reflection of the abundance of hosts found there, but also reflects the number of studies

performed, covering all of the oceans and many important seas around the globe, such as the Mediterranean, the Baltic and the Arctic (**Supplementary Table S2**). As expected, extreme environments, such as thermal springs (high temperatures) and hypersalines (high osmolarity), were those with the lowest viral diversity, with only 11 and four viral groups found in each, respectively. The families *Globuloviridae* and *Spiraviridae* were detected exclusively in thermal springs. The viruses of these families infect hyperthermophilic archaea, which are highly abundant in hot springs, thus explaining the exclusivity of those viruses in these environments. No viral group was exclusive to hypersaline environments. Curiously, viruses belonging to the families *Sphaerolipoviridae* and *Pleolipoviridae* (archaea-infecting viruses) have already been isolated and characterized from extreme environments (Luk et al., 2014); however, representatives of these groups were not detected by metaviromic approaches so far.

The absence of some viral groups in certain metaviromic studies might be due to the employed methodology, either in the sequencing platform/method and bioinformatic pipelines, in the type of genetic material that was analyzed (DNA or RNA), or even (and mainly) the procedures employed in the preparation of the samples for sequencing. The vast majority of studies target DNA viruses and use 0.2 μm porous filters during the processing of the collected samples (**Supplementary Table S2**). These strategies restrict the detection of a large part of the viruses (those with RNA genome) and also the giant DNA viruses (Halary et al., 2016), thus making a change in the protocols for the preparation of samples for metaviromic approaches necessary. Nevertheless, it is important to emphasize that the majority of the sequences found in metaviromic studies has no similarities with known sequences available from public databanks. This demonstrates that although the emergence of metagenomic techniques greatly contributed to the discovery of new viruses, even leading the ICTV executive committee to recently approve the use of such information for viral classification (Simmonds et al., 2017), the works on isolation and characterization, both genomically and biologically, should continue and be encouraged. With the association of biological/virological and metaviromic approaches, we might have new insights into the real diversity and distribution of viruses on Earth.

Human-Associated Viruses and Viral Tropism

Since human species is the one with more associated viruses officially recognized by the ICTV among all of the hosts analyzed here, the next step was to turn our attention to these viruses. Until recently, it was thought that about 200 viruses were associated with infections in humans, some with no direct evidence of causing any disease (Woolhouse et al., 2012). Here, we demonstrate that among the known virosphere, 320 virus species are related to human infections (**Supplementary Table S3**). Among them, 146 (45.6%) infect only humans; 116 (36.2%) infect humans and other mammals, some considered important zoonosis, such as rabies (*Rabies lyssavirus*), poxviruses (*Orthopoxvirus*), and hantaviruses (*Hantavirus*) (Shchelkunov, 2013; Jackson, 2016b; Jiang et al., 2017); and 58 (18.2%) are arboviruses (viruses transmitted by arthropods, including mosquitoes, sandflies and ticks) (**Figure 4A**). These viruses are classified within 26 families, wherein *Anelloviridae*, *Bunyaviridae*, and *Papillomaviridae* are the most significant, gathering 44% of the human viruses (**Figure 4B**). These viruses are highly variable, both structurally and genetically, using different replicative strategies. Although all groups of Baltimore's classification possess representatives of human viruses [except for viroids that infect only plants (Steger and Perreault, 2016)], the majority belong to groups I–V, with retroviruses accounting for less than 3% of viruses (**Supplementary Table S3**). Although they are the minority among human viruses, retroviruses were central to the emergence of mammals, thus also to humans, being pivotal components in placenta development (Chuong, 2013). In addition, the human immunodeficiency virus (HIV), the main representative of the group, is one the main life-threatening pathogens, being responsible for immunosuppressive conditions, paving the way to numerous severe secondary infections such as tuberculosis, systemic mycosis, Kaposi sarcoma, among others (Miceli et al., 2011; Godfrey-Faussett and Ayles, 2016; Govindan, 2016).

Many viruses are responsible for severe clinical manifestations, while others are related only to mild symptoms of disease or even asymptomatic infections. To have a better view of the tropism of human viruses and the most affected organic system, we built a network graph associating the viruses with different systems of the human body, according to clinical manifestations related to different viral infections. The viruses that have no direct evidence of causing disease were also included in the analysis. The integumentary, respiratory, and nervous systems were the main affected systems, with 92, 72, and 58 associated viruses, respectively (**Figure 4C**). The integumentary and respiratory systems are the most exposed to infection by different micro-organisms, since they are in direct contact with the environment, thus being expected to be the most affected by viruses. It is noteworthy that many viruses that affect the respiratory tract also affect the muscular system, a reflection of the viruses that cause only flu-like symptoms (**Supplementary Figure S3**). Unlike the two first systems, the nervous system is not directly exposed to the environment, thus making it curious that it is the third most frequently affected system by viruses. Since it is an extremely important and delicate system of the human body, several studies have been conducted to elucidate possible threats for its components, leading to the identification of a considerable range of viruses associated with diseases of the nervous systems. Many of these viruses are associated with severe cases of encephalitis and meningitis, such as herpesviruses (Granerod et al., 2010), lyssaviruses (Jackson, 2016a), and flaviviruses (Daep et al., 2014) (**Supplementary Table S4**), which is why they are target of intense investigation, to better understand the biology of these viruses, thus allowing the development of control mechanisms and possible treatments for diseases. Many of the viruses of the nervous system also affect others, mainly the respiratory and integumentary systems (**Supplementary Figure S3**). In that sense, some viruses are considerable pantropics, affecting

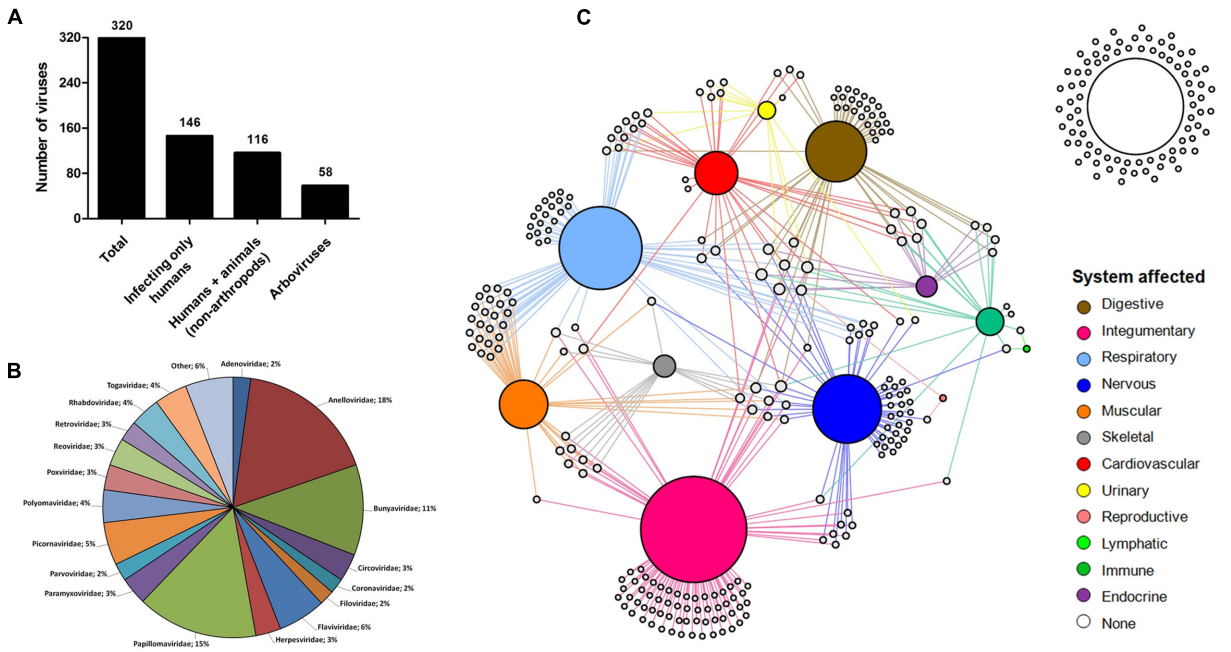


FIGURE 4 | Human viruses and affected systems. **(A)** Human-affecting viruses divided among infecting only humans, infecting humans and other mammals, and arboviruses. **(B)** Pie chart showing the classification of the viruses. A total of 27 groups are represented in the chart. Others: *Deltavirus, Hepadnaviridae, Hepeviridae, Caliciviridae, Picobirnaviridae, Pneumoviridae, Arenaviridae, Orthomyxoviridae, Astroviridae*. **(C)** Network graph showing the viral tropism. Each node represents a virus (white) and an organic system of the human body (colored nodes). The node diameter is proportional to the edge degree. The layout was generated using a force based algorithm followed by manual rearrangement to a better visualization of the connections.

different systems simultaneously, such as ebolavirus, dengue virus and rubella virus, affecting the cardiovascular (hemorrhagic fever), muscular (myalgia), skeletal (arthralgia), and nervous (encephalitis) systems, among others (**Supplementary Table S4**).

The reproductive and lymphatic systems are the least affected by viruses. The first is affected by only two viruses (mumps virus and Rio Bravo virus), responsible for cases of orchitis and oophoritis (Volkova et al., 2012). Although the herpesviruses and papillomaviruses are commonly associated with infections in the reproductive system, where they cause ulcerative lesions and warts in genital regions, we associated these viruses to the integumentary system, since their tropic site of infection is epidermal cells and not specific organs belonging to the reproductive tract. The lymphatic system has also only two associated virus species (*Human gammaherpesvirus 4* and *Primate T-lymphotropic virus 1*), both related to lymphoma cases. Although some viruses trigger lymph node inflammation, these are not considered the tropic site of infection for most viruses, so they are excluded from this analysis. It is possible that other viruses are related to these systems, as well as others included in this network, but further investigations are required. More studies are necessary regarding these systems, thus we can identify the viruses with tropism for these sites. Finally, 83 (26%) viruses analyzed in this work are not connected to any system since they are not related to any known disease so far (**Figure 4C**). The majority of these viruses belong to the family *Anelloviridae* (67.5%), which is mainly composed of the torque teno viruses. These viruses are present in most parts of people, as many metaviromic studies have demonstrated, but there is still no consensus that they carry any kind of loss for our health. As far as we know, they are part of the human virome along with many bacteriophages (Rascovan et al., 2016). Along with the anelloviruses, others have already been detected in human beings by metagenomic approaches, where the association with any disease remains under discussion, such as the giant mimiviruses and marseilleviruses (Popgeorgiev et al., 2013). While there is some evidence linking these viruses with human pathologies, we are still far from ending this debate.

CONCLUSION

It has been more than a century since the discovery of the first viruses. During this time, we have seen great advances in cellular and molecular biology and genetics, which have boosted achievements in the field of virology. Nevertheless, the results presented here show us that, even with great advances, we still know only a tiny fraction of the viral universe, mainly regarding the virus–host interaction. The discovery of giant viruses during the last decade was essential for us to realize how diverse and intriguing the virosphere is, triggering the search for new viruses in hosts completely ignored in the lens of virology. A break of concepts was established after those discoveries, taking us to think again what a virus is and what else is waiting to be discovered. Moreover, the advent of metaviromics had a unique contribution to the expansion of our knowledge about

the virosphere, mainly on the diversity and distribution of these microorganisms, but also with the discovery of new viruses (Alavandi and Poornima, 2012; Shi et al., 2016). However, we are still unable to define the host range of these new viruses with enough accuracy based only on genomic data. In that sense, the improvement of viral isolation techniques is important so that we can look deeper into how these new organisms interact with their hosts and the environment which they inhabit.

The analyses shown here provide a picture of what we know about the entire virosphere and their hosts, and confirm the anthropocentric view of the virology so far. It is likely that the network presented here (**Figure 2**) is largely more interconnected. However, further studies should be performed, especially searching for viruses in hosts that are not of primary human interest, such as environmental fungi and archaea, or even plants and animals that have no added medicinal or economic value. It is an arduous work, but with the improvement of viral isolation techniques and metaviromics, both fundamental tools to this task, it will be possible to continuously add new pieces to fulfill the virus–host network, providing a broader view of the viral universe. In that moment, possibly when science would once again be performed and applied to the understanding of the nature rather than serving the exclusive interests of human beings, we might see beyond just the tip of the iceberg.

AUTHOR CONTRIBUTIONS

RR, AA, and PB prepared the dataset. RR performed the analysis. RR wrote the manuscript. GT, EK, and JA designed the study. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by 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).

ACKNOWLEDGMENTS

We would like to thank our colleagues from Laboratório de Vírus of Universidade Federal de Minas Gerais. JA, GT, and EK are CNPq researchers. JA, EK, RR, and PB are members of a CAPES-COFECUB Project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01673/full#supplementary-material

FIGURE S1 | Amount of viruses associated by hosts (at genus level) separated by taxonomic group of the hosts. The total amount of hosts is depicted in the top of each column.

FIGURE S2 | The five hosts with more associated viruses for all six major taxonomic groups, evidencing that most of them is related to human interests. **(A)** Animalia, **(B)** Plantae, **(C)** Protist, **(D)** Fungi, **(E)** Bacteria, **(F)** Archaea. d, domesticated host; i, infection related host; b, biotechnology application host.

FIGURE S3 | Unipartite network graph showing the connections between organic systems according to the viruses that have tropism for more than one system. The nodes' diameter is proportional to the edge degree. The layout was generated using a force based algorithm followed by manual rearrangement to a better visualization of the connections. The thickness of the edges is proportional to the number of viruses that affect the two systems it connects.

TABLE S1 | Viruses and their hosts.

TABLE S2 | Technical information of metaviromic works.

TABLE S3 | Human-infecting viruses and other animals.

TABLE S4 | Tropism of human-infecting viruses and clinical manifestation.

REFERENCES

Alavandi, S. V., and Poornima, M. (2012). Viral metagenomics: a tool for virus discovery and diversity in aquaculture. *Indian J. Virol.* 23, 88–98. doi: 10.1007/s13337-012-0075-2

Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriol. Rev.* 35, 235–241.

Bastian, M., Heymann, S., and Jacomy, M. (2009). "Gephi: an open source software for exploring and manipulating networks," in *Third International AAAI Conference on Weblogs Social Media*, Paris, 361–362. doi: 10.1136/qshc.2004.010033

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (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. doi: 10.1073/pnas.0911354106

Cava, F., Hidalgo, A., and Berenguer, J. (2009). *Thermus thermophilus* as biological model. *Extremophiles* 13, 213–231. doi: 10.1007/s00792-009-0226-6

Chuong, E. B. (2013). Retroviruses facilitate the rapid evolution of the mammalian placenta. *Bioessays* 35, 853–861. doi: 10.1002/bies.201300059

Daep, C. A., Muñoz-Jordán, J. L., and Eugenin, E. A. (2014). Flaviviruses, an expanding threat in public health: focus on dengue, West Nile, and Japanese encephalitis virus. *J. Neurovirol.* 20, 539–560. doi: 10.1007/s13365-014-0285-z

Dietzgen, R. G., Mann, K. S., and Johnson, K. N. (2016). Plant virus-insect vector interactions: current and potential future research directions. *Viruses* 8, 1–21. doi: 10.3390/v8110303

Ďurkovič, J., Čaňová, I., Lagana, R., Kučerová, V., Moravčík, M., Priwitzer, T., et al. (2013). Leaf trait dissimilarities between Dutch elm hybrids with a contrasting tolerance to Dutch elm disease. *Ann. Bot.* 111, 215–227. doi: 10.1093/aob/mcs274

Ebrahimi-Mameghani, M., Sadeghi, Z., Abbasalizad Farhangi, M., Vaghef-Mehrabany, E., and Aliashrafi, S. (2016). Glucose homeostasis, insulin resistance and inflammatory biomarkers in patients with non-alcoholic fatty liver disease: beneficial effects of supplementation with microalgae *Chlorella vulgaris*: a double-blind placebo-controlled randomized clinical trial. *Clin. Nutr.* 36, 1001–1006. doi: 10.1016/j.clnu.2016.07.004

Enquist, L. W., and Racaniello, V. R. (2013). "Virology: from contagium fluidum to virome," in *Fields Virology*, eds D. M. Knipe and P. M. Howley (Philadelphia, PA: Lippincott Williams & Wilkins), 1–20.

Forterre, P. (2010). Giant viruses: conflicts in revisiting the virus concept. *Intervirology* 53, 362–378. doi: 10.1159/000312921

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., et al. (2016). Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* 166, 1397–1410.e16. doi: 10.1016/j.cell.2016.08.020

Godfrey-Faussett, P., and Ayles, H. (2016). Why are people living with HIV still dying of tuberculosis? *Lancet* 387, 1141–1143. doi: 10.1016/S0140-6736(16)00699-1

Goel, A. K. (2015). Anthrax: a disease of biowarfare and public health importance. *World J. Clin. Cases* 3, 20–33. doi: 10.12998/wjcc.v3.i1.20

Govindan, B. (2016). Recapitulation of acquired immuno deficiency syndrome associated Kaposi's sarcoma. *Indian J. Sex. Transm. Dis.* 37, 115–122. doi: 10.4103/0253-7184.192120

Granerod, J., Ambrose, H. E., Davies, N. W. S., Clewley, J. P., Walsh, A. L., Morgan, D., et al. (2010). Causes of encephalitis and differences in their clinical presentations in England: a multicentre, population-based prospective study. *Lancet Infect. Dis.* 10, 835–844. doi: 10.1016/S1473-3099(10)70222-X

Halary, S., Temmam, S., Raoult, D., and Desnues, C. (2016). Viral metagenomics: are we missing the giants? *Curr. Opin. Microbiol.* 31, 34–43. doi: 10.1016/j.mib.2016.01.005

Hillman, B. I., and Suzuki, N. (2004). Viruses of the chestnut blight fungus, *Cryphonectria parasitica*. *Adv. Virus Res.* 63, 423–472. doi: 10.1016/S0065-3527(04)63007-7

Holmfeldt, K., Solonenko, N., Shah, M., Corrier, K., Riemann, L., Verberkmoes, N. C., et al. (2013). Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12798–12803. doi: 10.1073/pnas.1305956110

International Committee on Taxonomy of Viruses - Taxonomy (2017). Available at: https://talk. ictvonline. org/taxonomy/w/ictv-taxonomy [accessed July 1, 2017].

Jackson, A. C. (2016a). Diabolical effects of rabies encephalitis. *J. Neurovirol.* 22, 8–13. doi: 10.1007/s13365-015-0351-1

Jackson, A. C. (2016b). Human rabies: a 2016 update. *Curr. Infect. Dis. Rep.* 18, 1–6. doi: 10.1007/s11908-016-0540-y

Jiang, H., Zheng, X., Wang, L., Du, H., Wang, P., and Bai, X. (2017). Hantavirus infection: a global zoonotic challenge. *Virol. Sin.* 32, 32–43. doi: 10.1007/s12250-016-3899-x

Kondo, H., Kanematsu, S., and Suzuki, N. (2013). Viruses of the white root rot fungus, *Rosellinia necatrix*. *Adv. Virus Res.* 86, 177–214. doi: 10.1016/B978-0-12-394315-6.00007-6

Korb, V. C., Chuturgoon, A. A., and Moodley, D. (2016). Mycobacterium tuberculosis: manipulator of protective immunity. *Int. J. Mol. Sci.* 17:131. doi: 10.3390/ijms17030131

Kristensen, D. M., Mushegian, A. R., Dolja, V. V., and Koonin, E. V. (2010). New dimensions of the virus world discovered through metagenomics. *Trends Microbiol.* 18, 11–19. doi: 10.1016/j.tim.2009.11.003

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., et al. (2003). A giant virus in amoebae. *Science* 299, 2033. doi: 10.1126/science.1081867

Langan, K. M., Kotsimbos, T., and Peleg, A. Y. (2015). Managing *Pseudomonas aeruginosa* respiratory infections in cystic fibrosis. *Curr. Opin. Infect. Dis.* 28, 547–556. doi: 10.1097/QCO.0000000000000217

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (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. doi: 10.1073/pnas.1320670111

Lipinska, A. P., Van Damme, E. J. M., and De Clerck, O. (2016). Molecular evolution of candidate male reproductive genes in the brown algal model Ectocarpus. *BMC Evol. Biol.* 16:5. doi: 10.1186/s12862-015-0577-9

Luk, A. W. S., Williams, T. J., Erdmann, S., Papke, R. T., and Cavicchioli, R. (2014). Viruses of haloarchaea. *Life (Basel)* 4, 681–715. doi: 10.3390/life4040681

Lwoff, A. (1957). The concept of virus the third marjory stephenson memorial lecture. *J. Gen. Microbiol.* 17, 239–253.

Menezes, C. B., Amanda Piccoli Frasson, A. P., and Tasca, T. (2016). Trichomoniasis – are we giving the deserved attention to the most common non-viral sexually transmitted disease worldwide? *Microb. Cell* 3, 404–418. doi: 10.15698/mic2016.09.526

Miceli, M. H., Díaz, J. A., and Lee, S. A. (2011). Emerging opportunistic yeast infections. *Lancet Infect. Dis.* 11, 142–151. doi: 10.1016/S1473-3099(10)70218-8

Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H., et al. (2016). Linking virus genomes with host taxonomy. *Viruses* 8:66. doi: 10.3390/v8030066

O'Malley, M. A. (2008). 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Philos. Biol. Biomed. Sci.* 39, 314–325. doi: 10.1016/j.shpsc.2008.06.005

Panahi, Y., Darvishi, B., Jowzi, N., Beiraghdar, F., and Sahebkar, A. (2016). *Chlorella vulgaris*: a multifunctional dietary supplement with diverse medicinal properties. *Curr. Pharm. Des.* 22, 164–173. doi: 10.2174/1381612822666151112145226

Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., et al. (2013). Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281–286. doi: 10.1126/science.1239181

Popgeorgiev, N., Temmam, S., Raoult, D., and Desnues, C. (2013). Describing the silent human virome with an emphasis on giant viruses. *Intervirology* 56, 395–412. doi: 10.1159/000354561

Raoult, D., and Forterre, P. (2008). Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* 6, 315–319. doi: 10.1038/nrmicro1858

Rascovan, N., Duraisamy, R., and Desnues, C. (2016). Metagenomics and the human virome in asymptomatic individuals. *Annu. Rev. Microbiol.* 70, 125–141. doi: 10.1146/annurev-micro-102215-095431

Reganold, J. P., and Wachter, J. M. (2016). Organic agriculture in the twenty-first century. *Nat. Plants* 2, 15221. doi: 10.1038/NPLANTS.2015.221

Shchelkunov, S. N. (2013). An increasing danger of zoonotic orthopoxvirus infections. *PLoS Pathog.* 9:e1003756. doi: 10.1371/journal.ppat.1003756

Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., et al. (2016). Redefining the invertebrate RNA virosphere. *Nature* doi: 10.1038/nature20167 [Epub ahead of print].

Sicard, D., and Legras, J. L. (2011). Bread, beer and wine: yeast domestication in the Saccharomyces sensu stricto complex. *C R Biol.* 334, 229–236. doi: 10.1016/j.crvi.2010.12.016

Siddiqui, R., and Khan, N. A. (2012). Biology and pathogenesis of Acanthamoeba. *Parasit. Vectors* 5:6. doi: 10.1186/1756-3305-5-6

Simmonds, P., Adams, M. J., Benkő, M., Breitbart, M., Brister, J. R., Carstens, E. B., et al. (2017). Consensus statement: virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* 15, 161–168. doi: 10.1038/nrmicro.2016.177

Steger, G., and Perreault, J. P. (2016). Structure and associated biological functions of viroids. *Adv. Virus Res.* 94, 141–172. doi: 10.1016/bs.aivir.2015.11.002

Suttle, C. A. (2005). Viruses in the sea. *Nature* 437, 356–361. doi: 10.1038/nature04160

Thornton, P. K. (2010). Livestock production: recent trends, future prospects. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 365, 2853–2867. doi: 10.1098/rstb.2010.0134

Thurber, R. V. (2009). Current insights into phage biodiversity and biogeography. *Curr. Opin. Microbiol.* 12, 582–587. doi: 10.1016/j.mib.2009.08.008

Vainio, E. J., and Hantula, J. (2015). Taxonomy, biogeography and importance of Heterobasidion viruses. *Virus Res.* 219, 2–10. doi: 10.1016/j.virusres.2015.10.014

Vila, J., Sáez-López, E., Johnson, J. R., Römling, U., Dobrindt, U., Cantón, R., et al. (2016). *Escherichia coli*: an old friend with new tidings. *FEMS Microbiol. Rev.* 40, 437–463. doi: 10.1093/femsre/fuw005

Volkova, E., Tesh, R. B., Monath, T. P., and Vasilakis, N. (2012). Full genomic sequence of the prototype strain (M64) of Rio Bravo virus. *J. Virol.* 86, 4715. doi: 10.1128/JVI.00331-12

Woese, C. R. (2002). On the evolution of cells. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8742–8747. doi: 10.1073/pnas.132266999

Woolhouse, M., Scott, F., Hudson, Z., Howey, R., and Chase-Topping, M. (2012). Human viruses: discovery and emergence. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 2864–2871. doi: 10.1098/rstb.2011.0354

Zhang, C., Krause, D. J., and Whitaker, R. J. (2013). *Sulfolobus islandicus*: a model system for evolutionary genomics. *Biochem. Soc. Trans.* 41, 458–462. doi: 10.1042/BST20120338

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Human Microbiome Journal 11 (2019) 100050

Contents lists available at ScienceDirect

Human Microbiome Journal

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

Virus and microbiota relationships in humans and other mammals: An evolutionary view

Maurício Teixeira Lima, Ana Cláudia dos Santos Pereira Andrade, Graziele Pereira Oliveira, Jacques Robert Nicoli, Flaviano dos Santos Martins, Erna Geessien Kroon, Jônatas Santos Abrahão*

Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais 31270-901, Brazil

ARTICLE INFO

Keywords:
Indigenous microbiota
Virus-host interactions
Virus-microbiota interactions
Microbiota-host interactions
Evolution

ABSTRACT

In the last decades, studies have revealed multiple and strong correlations between the host and its commensal microbiota consisting of bacteria, protozoa, fungi and viruses. This associated microbiota can positively or negatively influence the course of a wide range of infections. Here, we review the interactions between the host and its viral microbiota and discuss new paradigms from an evolutionary perspective. The viral adaptation to a microbial environment in a co-evolutionary approach is highlighted, as well as viral cross transmission in the context of the barriers imposed by the indigenous microbiota. In addition to reviewing the host-microbiota-virus relationships, we focus the discussion on microbiota-virus interactions that could be applied to preventive and therapeutic treatments.

1. Introduction

Viruses are the most abundant biological entities on Earth and have evolved with prokaryotes and eukaryotes for thousands of years. The abundance of viruses varies according to the environment and sometimes is relative to bacterial activity and colonization [1]. The human gut harbors a dense and complex microbial ecosystem, with presents not only prokaryotic and eukaryotic organisms but also viruses (virobiota, and their genes – virome). This indigenous microbiota can be associated with the host cells (eukaryotic virobiota) or with some of the approximately 2776 prokaryotic species that inhabit it (prokaryotic virobiota) and this interation may be beneficial or detrimental to the host [2]. An example of a positive effect is the adherence of phages to mucus forming an antimicrobial barrier in various host mucosal surfaces [3]. This co-evolutionary mechanism is called “non-host-derived immunity” and acts primarily controlling the abundance and equilibrium of bacterial populations [3,4]. The evolutionary battle between viruses and prokaryotes was reported in the last years by studies on the “Kill the Winner” hypothesis, horizontal genetic exchange, CRISPR-encoding bacteria and viral anti-CRISPR proteins [5–7]. Investigation of these relationships has provided important biotechnological tools that can be used for genetic engineering such as CRISPR-Cas9 genome editing human cells [6].

Recent studies have shown that the host’s normal microbiota is able to influence the infections caused by various families of animal viruses [8–10]. Microbiota-virus interactions have been studied in germ-free mice or antibiotic-treated mice models highlighting the opposing modulating effects of commensal bacteria on the course of viral infections [9,10]. Commensal bacteria can potentially influence viral infections either hindering or promoting the viral infection and sometimes aggravate the disease [8,9]. Bacteria commonly isolates from human nasopharynx as *Staphylococcus aureus*, *Pseudomonas* species, *Streptococcus pneumonia*, *Haemophilus influenzae* and *Streptococcus pyogenes* has been associated with increased risk of death in adults and children infected with influenza [11].

This review highlights the important role of microbiota-virus interactions through an evolutionary perspective, emphasizing the viral adaptations to the microbial environment and the use of available resources by viruses. The cooperation or the competition with other components of the indigenous microbiota, as well as the co-evolution with host and viral cross-species transmission in the context of the barriers imposed by the endogenous microbiota are also addressed.

2. Viruses versus microbial ecosystem

Viral particles face numerous host-related challenges to reach the

* Corresponding author at: Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil.
E-mail address: jonatas.abrahao@gmail.com (J.S. Abrahão).

https://doi.org/10.1016/j.humic.2018.11.001

Available online 24 January 2019

permissive cells, such as tissue specificities, body temperature and epithelial secretions including IgA, defensins and a mucus barrier, as well as environmental modifications due to microbiota metabolism and cellular composition such as pH, redox potential, lipopolysaccharide (LPS) and glycans. The gastrointestinal microbiota is the most complex and diverse ecosystem in mammals, quite different when compared to those present in other body sites, and there is a considerable variation in the constituents of the gut microbiota among apparently healthy individuals [12].

The presence of the microbiota or its products is associated with increases in the viral fitness for all enteric viruses studied so far, including *Enterovirus C* (poliovirus) [13,14], *Mammalian reovirus* [13], *Rotavirus* A [15], *Norwalk virus* (norovirus) [16–19] and *Mouse mammary tumor virus* (MMTV) [20,21] (an enteric retrovirus). In this context, some findings suggest different mechanisms by which the enteric viruses could use bacteria and their products to withstand environmental adversities and cross the host cell barriers.

A study of poliovirus was the first to show that viral exposure to bacteria enhanced host cell binding and infection by the virus [13]. The enhancement of viral infectivity did not require live bacteria, and the presence of bacterial surface polysaccharides, including LPS and peptidoglycan (PGN), was sufficient [13]. LPS is the major cell wall component of Gram-negative bacteria with highest concentrations in the gut lumen [22]. Poliovirus can use LPS to promote attachment to the surface of permissive cells through direct facilitation of viral binding to its poliovirus receptor (Fig. 1A). In addition, LPS can enhance virion environmental stability by increasing its thermostability and resistance to chlorine bleach [13,14]. A specific residue in the capsid protein of poliovirus VP1 was shown to be crucial for stabilization, and this ability is important to prevent premature conformational changes before uncoating [14] (Fig. 1B). A mechanism similar to the LPS-mediated stimulation of poliovirus was observed for human norovirus when it was discovered that it could infect human B cells [16]. Some specific commensal bacteria express a glycan called histo-blood group antigen (HBGA) that correlated with the ability of norovirus to attach and infect B cells [18]. The isolated HBGA was sufficient to stimulate viral attachment to the surface of B cells [19] (Fig. 1C) by using a mechanism apparently very similar to that of poliovirus-LPS attachment to its host receptor. However, the receptor used by human norovirus remains unknown precluding an understanding of the mechanism by which bacterial HBGA stimulates viral attachment [19].

It is not yet clear how viruses cross all of the barriers to reach the host cells, such as those which separate virus from B cells and the human norovirus-HBGA complex located in the intestinal lumen. Some opportunistic members of the commensal microbiota contribute to the process of injury of the gut barriers by primarily destroying the mucus layer or even the enterocytes by using the enzymatic apparatus [23]; viral particles could use these passages to reach the target cells. Furthermore, many bacteria can translocate, which is the ability to pass through the intestinal epithelium from the lumen to the internal compartment [24].Viruses may adhere to bacterial surfaces during translocation, which occurs through the transcellular (inside enterocytes) or the paracellular (by passing through the intercellular space between the cells) pathways. This phenomenon described here as “microbial phoresy” is defined in ecology as an inter-species biological interaction where an organism is mechanically transported by its host. Phoresy is frequently used within the animal kingdom and is a type of commensalism where neither organism is physiologically dependent on the other. Recently, the binding of human norovirus around the outer cell surfaces and pili structures of bacteria was described, but without apparent localization [25] (Fig. 1D). A similar function may be present in other viruses of different viral families seen the affinities between these and bacterial surface components [26]. In conclusion, viruses can interact with the microbiota and their products to increase viral fitness through virion stability and enhanced binding to the surface of target host cells. (Fig. 1A–D).

3. Host immune environment: From tolerance to battlefield

The intestinal microbiota interacts with the whole host immune system and consequently with antiviral immune responses. However some viruses have evolved to use this phenomenon, the microbiota being directly involved in viral evasion of the host immune system and in induction of a tolerogenic microenvironment. The establishment of a tolerogenic microenvironment is made by specific regulatory T cells, highly prevalent in the intestine and responsible for recognition of commensal microbiota and for the immunological tolerance to many of its non-pathogenic components [27].Viruses may use regulatory T cells to influence antiviral immune responses, such as for the persistence of MMTV retrovirus infection in mice pups (Fig. 2). MMTV seems to bind directly to LPS and then incorporates LPS-binding host proteins into its envelope, such as CD14, MD2 and TLR4. Virions isolated from knockout mice for LPS-binding proteins were unable to bind to LPS [20,21]. It is important to highlight the immunostimulatory nature of LPS in generating sequential events in which binding of LPS to TLR4 drives the production of IL-6, which then induces IL-10 secretion [20,21] (Fig. 2). In addition, virion-LPS binding is essential for viral transmission in pups that ingested MMTV-laden maternal milk [20]. Furthermore, infected mouse pups failed to produce detectable antiviral antibodies when immunized orally with viral antigens, whereas those which were exposed to MMTV intraperitoneally were not tolerant to MMTV antigens [20]. These data revealed that interactions between the microbiota and MMTV promotes tolerance to viral antigens and facilitates the establishment of a persistent viral infection.

Another virus that has been shown to benefit from the bacterial modulation of the host immune system is murine norovirus (MNV), which presents a bacteria-mediated persistence [28]. In vivo assays using mice treated with antibiotics for bacterial depletion showed that the treatment inhibits persistent MNV infection in the intestine [28]. The tissue infection and systemic viral replication were not affected by antibiotic's treatment and the persistent infection occurred only when the microbiota was recomposed [28]. Additionally, mice lacking IFN-λ receptor were persistently infected with a MNV irrespective of the presence of commensal microorganisms [28,29]. The hypothesis currently accepted is that commensal bacteria suppress the production of IFN-λ upon their interaction with MNV [28,29]. Similarly, this cytokine also controls rotavirus infection in murine model [30]. However, a study showed that MNV can replace the beneficial function of microbiota in germ-free or antibiotic-treated mice restored intestinal morphology and lymphocyte function without inducing overt inflammation and disease [31]. The IFN-α receptor was associated with the ability of MNV to compensate for bacterial depletion and kept the intestinal homeostasis [31].

As pointed out above, the modulation of the immune system by the microbiota can promote viral infection, but can inversely benefit the host depending on the viral agent. The presence of the microbiota is essential for an effective immune response against the *Vaccinia virus* (VACV), a large and complex enveloped virus belonging to the *Poxviridae* family [32]. An *in vivo* study demonstrated that VACV presented a similar profile of systemic infection in germ-free and immunosuppressed mice models, whereas conventional mice were refractory to this infection [32]. Additional *in vivo* studies have also demonstrated a microbiota-mediated protection against the *Influenza A virus,* in which the microbiota activated the inflammasome [33]. The inflammasome activation induced migration of dendritic cells from the lung to the draining lymph node, to stimulate influenza-specific T-cell responses (Fig. 3A). Antibiotic treatment in mice caused microbiota depletion and increased the animals’ susceptibility to the virus [34]. Additionally, recent studies highlighted the importance of signals derived from commensal bacteria that can calibrate the activation threshold of innate immunity, and revealed an interplay between commensal and antiviral interferon signaling pathways in macrophages, involved in responses to *Influenza A virus* [35]. Conversely,

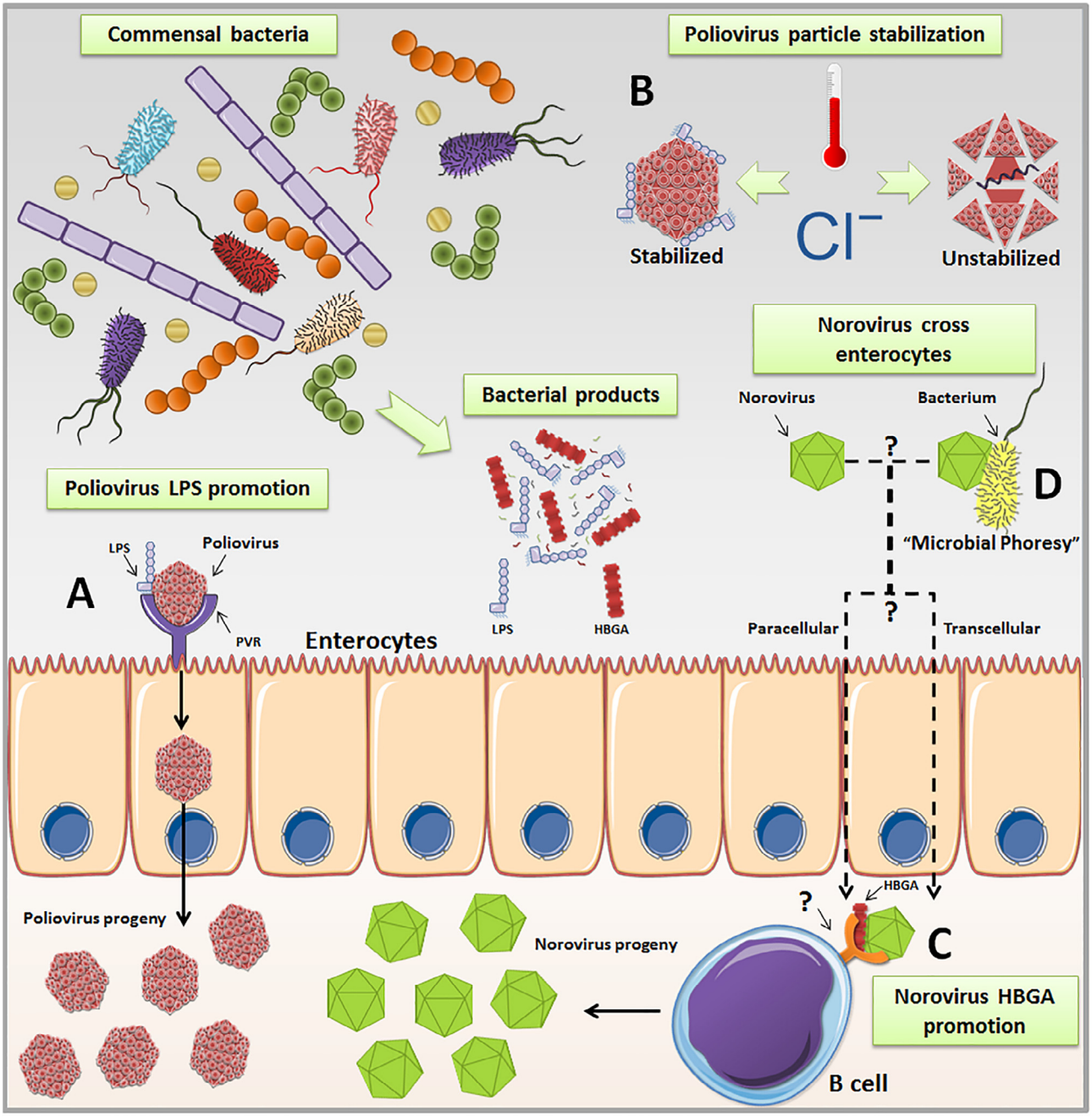


Fig. 1. Representative mechanisms by which bacteria and their products are used by viruses to cross the barriers to the host's cells and increase viral fitness. (A) Viral exposure to bacterial products enhanced binding to the host cell and infection caused by poliovirus. (B) Association of poliovirus particles with LPS increases their thermostability and resistance to inactivation by chlorine bleach. (C) HBGA-mediated stimulation mechanism used by norovirus to infect human B cells. (D) A commensal bacterium contributing to the ability of norovirus to cross all of the barriers to the host cells by the "microbial phoresy" mechanism.

studies also showed that respiratory influenza infection could cause digestive diseases such as intestinal immune injury and secondary infections in gut [36,37]. The intestinal immune injury may have resulted from an altered intestinal microbiota composition mediated by IFN-γ produced by lung-derived T-cells and recruited into the intestine (Fig. 3B) [36].

Therefore, in the host immune system environment, the cross-talk between microbiota and virus shows a two-way pattern involving regulator and regulated agents, both the microbiome and the viral infection having had alterations in their composition and fitness, respectively. In summary, the triangular relationships between viruses, microbiota and immune system are complex and vary from tolerance to dysbiosis in a phenomenon shaped by co-evolution.

4. From co-evolution to cross species transmission

In host-microbiota-virus co-evolution, each element of this relationship exerts selective pressures the others, thereby affecting the evolutionary aspects in general. Many organisms have been described for their probiotic effect on various types of viral infections, leading to consequent applications in preventive and therapeutic approaches [38–41]. Specific members of the microbiota may have a promoter or inhibitory effect on viral infections in the original or new host.

The human female vaginal microbiota co-evolved with the entire genus *Lactobacillus* [42]. These bacteria responsible for production of antiviral agents such as hydrogen peroxide (H_2O_2) and lactic acid are dominant in the vaginal ecosystem. The hydrogen peroxide produced

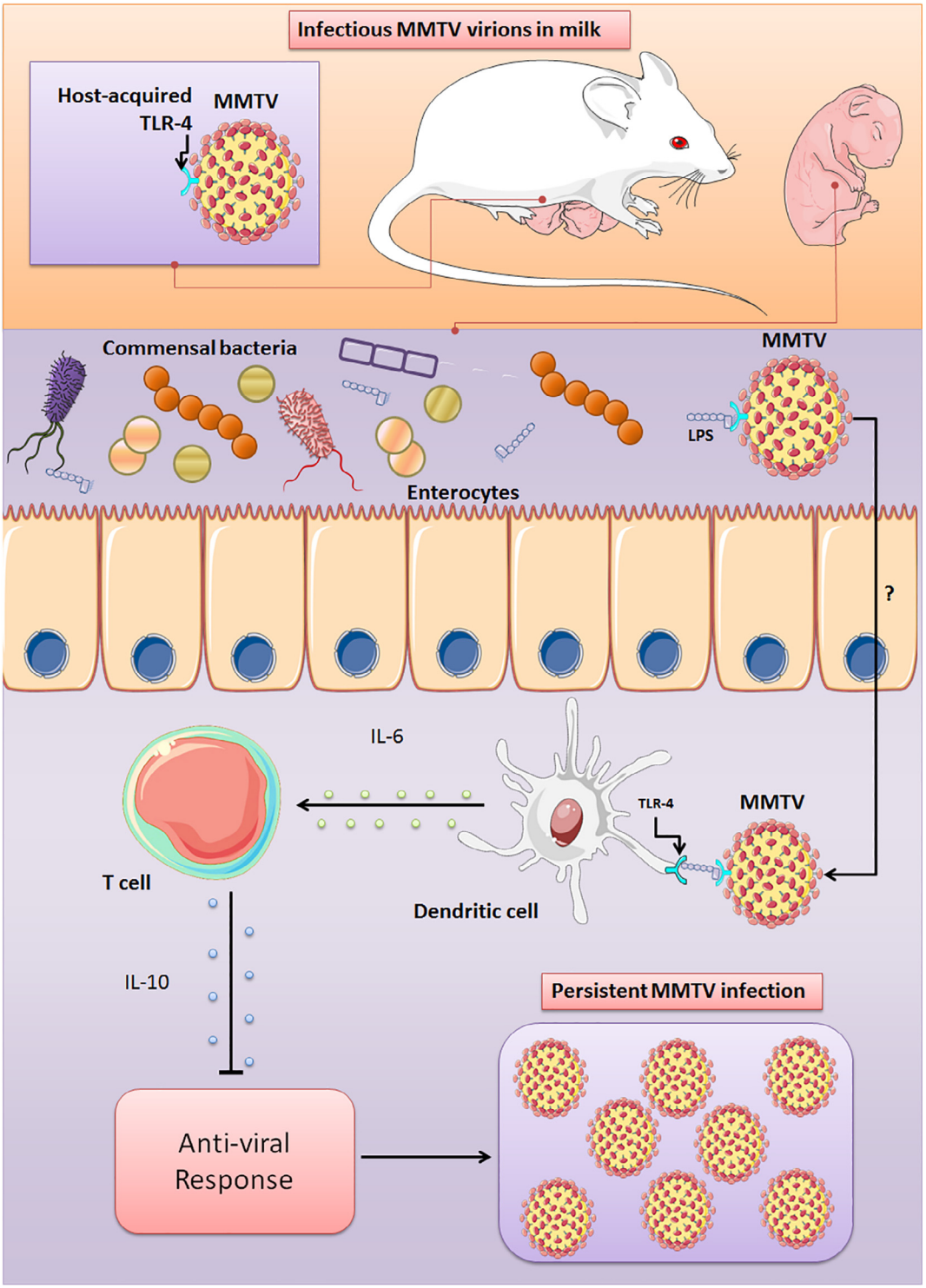


Fig. 2. Establishment of persistent MMTV infection in mice pups wherein regulatory T cells influence the antiviral immune response. The engagement of TLR4 (that was incorporated by MMTV) on LPS drives the production of IL-6, which induces IL-10 secretion. In this immunosuppressive microenvironment, MMTV is able to establish a persistent infection.

by *Lactobacillus* plays an important role as a natural antimicrobial agent in the vaginal ecosystem and is toxic to a numerous organisms, including viruses such as *Human immunodeficiency virus 1* (HIV-1) [43] and *Human alphaherpesvirus 2* (HSV-2) [44]. Another important *Lactobacillus* antimicrobial product is lactic acid, which is responsible for homeostasis of the female vaginal pH ($\leq$4.5)[45].This pH is the lowest in the vaginal ecosystem among all mammals and differs greatly from other primates ($\sim$pH 7.0) [45]. Lactobacilli generate L-lactic acid as a final product of the glycogen metabolism produced by mucosal epithelial cells. This acidic environment can inhibit the growth of several potentially pathogenic species, such as *Chlamydia trachomatis*, *Gardnerella vaginalis* in addition to inactivating HIV, HSV-2 and human papillomavirus type 16 (HPV-16), an *Alphapapillomavirus 9* [46–53] (Fig. 4A). The E5 protein of HPV-16, responsible for viral transformation, is known to be particularly susceptible to low pH [53]. In addition, vaginal microbiota dominated by *Lactobacillus gasseri* was associated

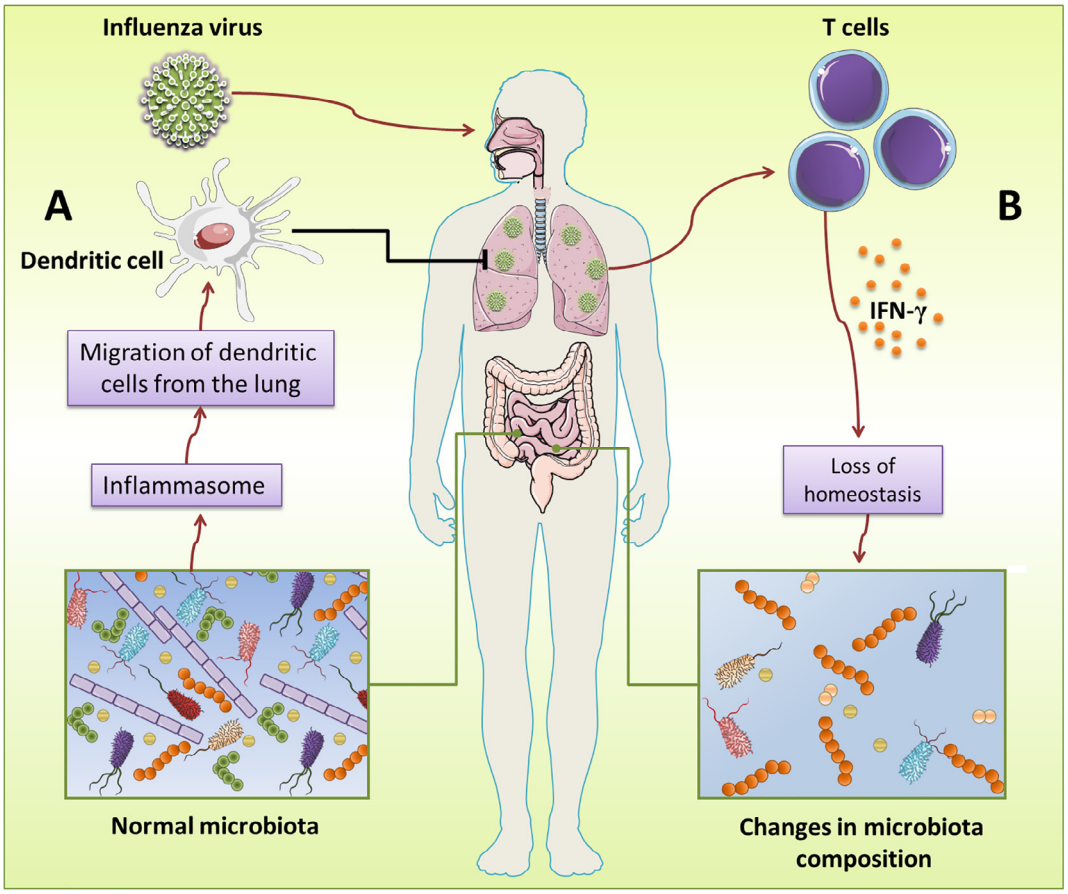


Fig. 3. Representative mechanism of the influence of the microbiota on influenza infection. (A) The migration of dendritic cells from the lung induced by inflammasome activation acts on influenza-specific T-cell responses characterizing a mechanism of microbiota-mediated protection against the influenza. (B) Influenza infection changes the intestinal microbiota composition mediated by IFN-γ produced by lung-derived T-cells recruited into the intestine.

with a faster clearance rate of detectable human papillomavirus (HPV) [54].

A comparative study of the primate vaginal microbiome showed that *Lactobacillus* dominance is an exclusive characteristic of humans and does not occur in non-human primates [45]. Importantly, women with bacterial vaginosis exhibit compositional similarities with non-human primates and a higher vaginal pH (pH = 5.5) than healthy women [55] (Fig. 4B). Bacterial vaginosis increases HSV-2 infection and has been linked to an increased risk of HIV-1 acquisition [56–58]. Vaginal microbiota may influence women's risk of HIV acquisition by some groups of microorganisms promotegenital inflammation and produce HIV-inducing factors in vaginal fluid, sialidases and mucinases that disrupt the protective cervicovaginal mucus layer [58–61]. Differences in the vaginal microbial diversity and concentrations of *Parvimonas* species types 1 and 2, *Gemella asaccharolytica*, *Mycoplasma hominis*, *Leptotrichia/Sneathia*, *Eggerthella* species type 1 and *Megasphaera* were significantly associated with increased risk of HIV acquisition in african women [58]. The majority of new HIV infections in Africa reached women women, unlike other parts of the world. Furthermore vaginal dysbiosis accounts for 20–30% of the population attributable risk of HIV acquisition in african women [58]. Additionally, in a study with Costa Rican pre-menopausal women, vaginal pH greater than 5.0 was shown to be significantly associated with a 10–20% increased risk of HPV positivity [62]. The characterization of the vaginal bacterial microbiota associated with HIV, HSV-2 and HPV risk provide important targets for future prevention research.

From an evolutionary perspective, humans and non-human primates differ considerably in mating habits, diet, estrus cycles, sexual behavior, gestation period and vaginal pH, these factors being associated with differences in microbial composition. To cross between animal species, a virus must overcome the barriers imposed by host biology and microbiota (Fig. 4C). An event related to this fact occurred recently for the HIV-1 origin through multiple events of cross-species transmission of *Simian immunodeficiency virus* (SIV) [63] (Fig. 4). In this case, it is essential to consider the aspects of the barriers imposed by the microbiota to better understand the path of the virus until it reaches its pandemic form and its epidemiology. This situation relates to the probability of women becoming infected with HIV-1 via vaginal intercourse, which is significantly lower than that of rectal or parenteral transmission [64].

5. Conclusion

In the last years, studies with poliovirus, norovirus, MMTV and influenza virus showed strict relationships between the microbiota and virus infections. These studies raised new questions and hypotheses that have remained largely unanswered, including the mechanism by which some viruses pass through the intestinal mucosa to reach their target cells. In view of the current findings about the affinity of viruses with the bacterial surface, the "microbial phoresy" traces a microbial parallel with the macroscopic animal world that could be an interesting model to explain how the viruses travel through the intestinal microbiota.

Another very interesting issue discussed here was the unusual features of the human healthy vaginal microbiota, such as the low pH and H_2O_2 presence. These features are effective barriers against various pathogens of the vaginal tract and a potential barrier against viral species cross transmission, such as for SIV-HIV where the probability of infection by this route is much lower than by others, but increased in

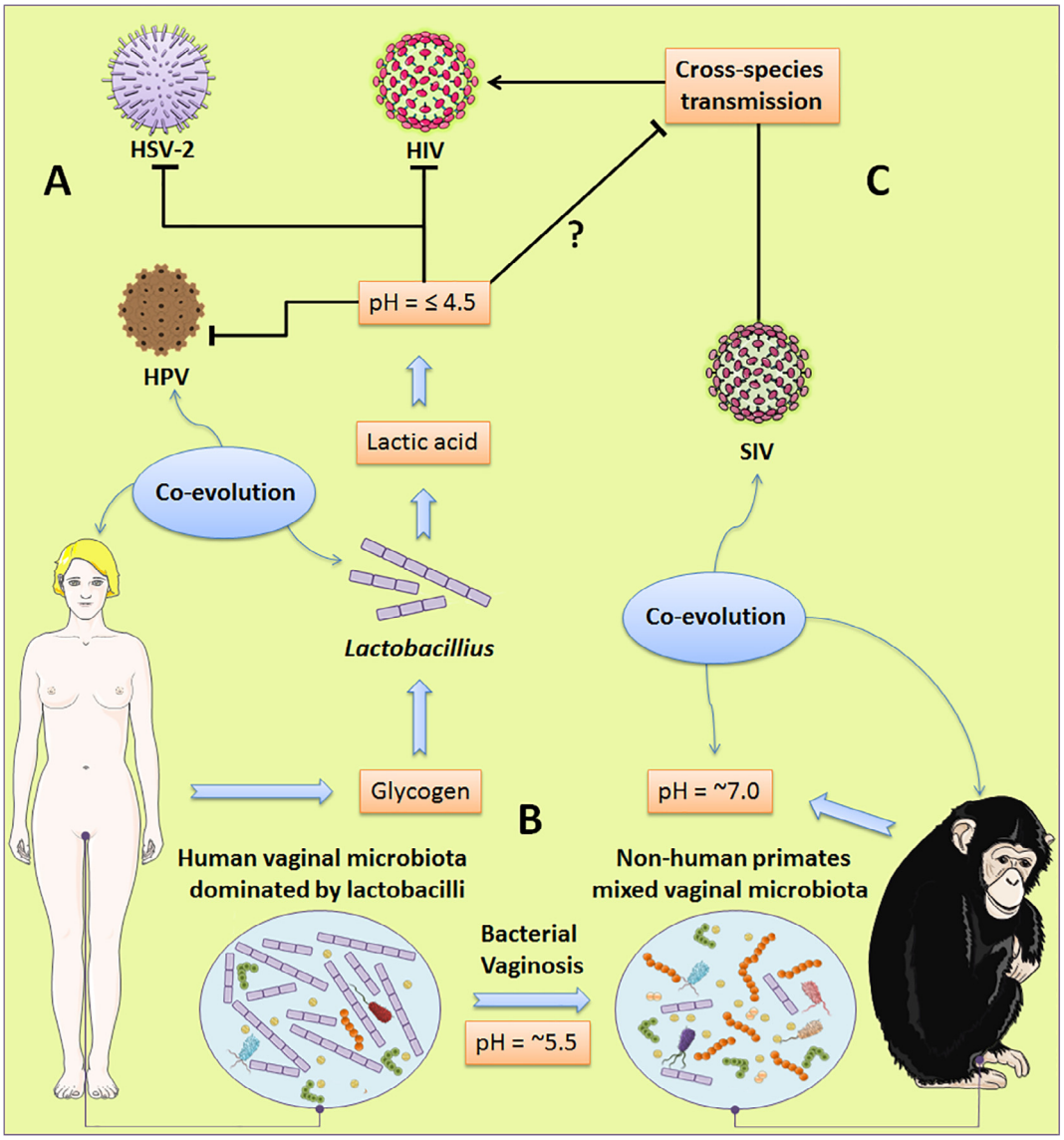


Fig. 4. The co-evolution of the human female vaginal microbiota with the *Lactobacillus* genus can inhibit several pathogenic species and create potential barriers to cross-species transmission. (A) An acidic environment can inhibit HIV, HSV-2 and HPV. (B) *Lactobacillus* dominance profile is exclusive to humans and does not occur in non-human primates. Women with bacterial vaginosis exhibit a higher vaginal pH than healthy women. (C) The origin of HIV-1 from cross-species transmission of *Simian immunodeficiency virus* (SIV). An acidic environment can create a potential barrier to cross-species transmission.

the situation of vaginal dysbiosis. An evolutionary hypothesis, named "disease risk hypothesis," proposes that humans face higher sexually transmitted disease risk than non-human mammals [65,66]. Humans show a more continuous sexual receptivity throughout the menstrual cycle, pregnancy, and the post-partum period [67]. This fact generates a selective pressure for the need of protective mechanisms, which could explain why *Lactobacillus* populations are higher in the human vaginal ecosystem compared to those of mammals with less frequent sexual contact [55].

We emphasize the need for further studies to better understand the regulatory relationships between the microbiota, the viral agents and the host immune system. The use of genomic and metagenomic tools is fundamental for the advancement of studies on microbiota-virus-host relationships given that little has been described about the amplitude of the virosphere. Taking as inspiration the famous phrase cited by the evolutionary biologist Theodosius Dobzhansky who says "Nothing in Biology Makes Sense Except in the Light of Evolution," we believe that an evolutionary perspective on microbiota-virus relationships is essential to discuss and unveil new aspects that remain in the dark.

Declaration of interest

We have no conflict of interest to declare.

Author contributions

Wrote the paper: MTL, ACSPA and GPO; designed the study, revised the content and final approved the final version: MTL, ACSPA, GPO, JRN, FSM, EGK and JSA.

Acknowledgements

We are grateful to our colleagues from Laboratório de Vírus of Universidade Federal de Minas Gerais. In addition, this work was supported by the grants Pro-Reitoria de Pesquisa da Universidade Federal de Minas Gerais (PRPq-UFMG) and by the support agencies 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).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humic.2018.11.001.

References

[1] Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, et al. Uncovering Earth's virome. Nature 2016;536:425–30. https://doi.org/10.1038/nature19094.

[2] Bilen M, Dufour J-C, Lagier J-C, Cadoret F, Daoud Z, Dubourg G, et al. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. Microbiome 2018. https://doi.org/10.1186/s40168-018-0485-5.

[3] De Paepe M, Leclerc M, Tinsley CR, Petit M-A. Bacteriophages: an underestimated role in human and animal health? Front Cell Infect Microbiol 2014;4:39. https://doi.org/10.3389/fcimb.2014.00039.

[4] Barr JJ, Auro R, Furlan M, Whiteson KL, Erb ML, Pogliano J, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc Natl Acad Sci 2013. https://doi.org/10.1073/pnas.1305923110.

[5] Maslov S, Sneppen K. Population cycles and species diversity in dynamic Kill-the-Winner model of microbial ecosystems. Sci Rep 2017. https://doi.org/10.1038/srep39642.

[6] Stanley SY, Maxwell KL. Phage-Encoded Anti-CRISPR Defenses. Annu Rev Genet 2018. https://doi.org/10.1146/annurev-genet-120417-031321.

[7] Hatfull GF, Hendrix RW. Bacteriophages and their genomes. Curr Opin Virol 2011. https://doi.org/10.1016/j.coviro.2011.06.009.

[8] Wilks J, Golovkina T. Influence of microbiota on viral infections. PLoS Pathog 2012;8. https://doi.org/10.1371/journal.ppat.1002681.

[9] Wilks J, Beilinson H, Golovkina TV. Dual role of commensal bacteria in viral infections. Immunol Rev 2013;255:222–9. https://doi.org/10.1111/imr.12097.

[10] Karst SM. The influence of commensal bacteria on infection with enteric viruses. Nat Rev Microbiol 2016;14:197–204. https://doi.org/10.1038/nrmicro.2015.25.

[11] Chertow DS, Memoli MJ. Bacterial coinfection in influenza: a grand rounds review. JAMA - J Am Med Assoc 2013. https://doi.org/10.1001/jama.2012.194139.

[12] Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol 2015;31:69–75. https://doi.org/10.1097/MOG.0000000000000139.

[13] Kuss SK, Best GT, Etheredge CA, Pruijssers AJ, Frierson JM, Hooper LV, et al. Intestinal microbiota promote enteric virus replication and systemic pathogenesis. Science (80-) 2011;334:249–52. https://doi.org/10.1126/science.1211057.

[14] Robinson CM, Jesudhasan PR, Pfeiffer JK. Bacterial lipopolysaccharide binding enhances virion stability and promotes environmental fitness of an enteric virus. Cell Host Microbe 2014;15:36–46. https://doi.org/10.1016/j.chom.2013.12.004.

[15] Varyukhina S, Freitas M, Bardin S, Robillard E, Tavan E, Sapin C, et al. Glycan-modifying bacteria-derived soluble factors from Bacteroides thetaiotaomicron and Lactobacillus casei inhibit rotavirus infection in human intestinal cells. Microbes Infect 2012;14:273–8. https://doi.org/10.1016/j.micinf.2011.10.007.

[16] Jones MK, Watanabe M, Zhu S, Graves CL, Keyes LR, Grau KR, et al. Enteric bacteria promote human and mouse norovirus infection of B cells. Science (80-) 2014. https://doi.org/10.1126/science.1257147.

[17] Tan M, Jiang X. Norovirus and its histo-blood group antigen receptors: An answer to a historical puzzle. Trends Microbiol 2005;13:285–93. https://doi.org/10.1016/j.tim.2005.04.004.

[18] Miura T, Sano D, Suenaga A, Yoshimura T, Fuzawa M, Nakagomi T, et al. Histo-blood group antigen-like substances of human enteric bacteria as specific adsorbents for human noroviruses. J Virol 2013;87:9441–51. https://doi.org/10.1128/JVI.01060-13.

[19] Karst SM. Identification of a novel cellular target and a co-factor for norovirus infection – B cells & commensal bacteria. Gut Microbes 2015;6:266–71. https://doi.org/10.1080/19490976.2015.1052211.

[20] Kane M, Case LK, Kopaskie K, Kozlova A, MacDearmid C, Chervonsky AV, et al. Successful transmission of a retrovirus depends on the commensal microbiota. Science (80-) 2011;334:245–9. https://doi.org/10.1126/science.1210718.

[21] Wilks J, Lien E, Jacobson AN, Fischbach MA, Qureshi N, Chervonsky AV, et al. Mammalian lipopolysaccharide receptors incorporated into the retroviral envelope augment virus transmission. Cell Host Microbe 2015;18:456–62. https://doi.org/10.1016/j.chom.2015.09.005.

[22] Guo S, Al-Sadi R, Said HM, Ma TY. Lipopolysaccharide causes an increase in intestinal tight junction permeability in vitro and in vivo by inducing enterocyte membrane expression and localization of TLR-4 and CD14. Am J Pathol 2013;182:375–87. https://doi.org/10.1016/j.ajpath.2012.10.014.

[23] Qin X, Sheth SU, Sharpe SM, Dong W, Lu Q, Xu D, et al. The mucus layer is critical in protecting against ischemia-reperfusion-mediated gut injury and in the restitution of gut barrier function. Shock 2011;35:275–81. https://doi.org/10.1097/SHK.0b013e3181f6aaf1.

[24] Fouts DE, Torralba M, Nelson KE, Brenner DA, Schnabl B. Bacterial translocation and changes in the intestinal microbiome in mouse models of liver disease. J Hepatol 2012;56:1283–92. https://doi.org/10.1016/j.jhep.2012.01.019.

[25] Almand EA, Moore MD, Outlaw J, Jaykus LA. Human norovirus binding to select bacteria representative of the human gut microbiota. PLoS ONE 2017(12). https://doi.org/10.1371/journal.pone.0173124.

[26] Rodrigues RAL, dos Santos Silva LK, Dornas FP, de Oliveira DB, Magalhães TFF, Santos DA, et al. Mimivirus fibrils are important for viral attachment to the microbial world by a diverse glycoside interaction repertoire. J Virol 2015. https://doi.org/10.1128/JVI.01976-15.

[27] Wells JM, Rossi O, Meijerink M, van Baarlen P. Epithelial crosstalk at the microbiota-mucosal interface. Proc Natl Acad Sci USA 2011;108(Suppl):4607–14. https://doi.org/10.1073/pnas.1000092107.

[28] Baldridge MT, Nice TJ, McCune BT, Yokoyama CC, Kambal A, Wheadon M, et al. Commensal microbes and interferon-determine persistence of enteric murine norovirus infection. Science (80-) 2015;347:266–9. https://doi.org/10.1126/science.1258025.

[29] Nice TJ, Baldridge MT, McCune BT, Norman JM, Lazear HM, Artyomov M, et al. Interferon-λ cures persistent murine norovirus infection in the absence of adaptive immunity. Science 2015;347:269–73. https://doi.org/10.1126/science.1258100.

[30] Pott J, Mahlakoiv T, Mordstein M, Duerr CU, Michiels T, Stockinger S, et al. IFN-lambda determines the intestinal epithelial antiviral host defense. Proc Natl Acad Sci USA 2011;108:7944–9. https://doi.org/10.1073/pnas.1100552108.

[31] Kernbauer E, Ding Y, Cadwell K. An enteric virus can replace the beneficial function of commensal bacteria. Nature 2014. https://doi.org/10.1038/nature13960.

[32] Lima MT, Andrade ACSP, Oliveira GP, Calixto RS, Oliveira DB, Souza ÉLS, et al. Microbiota is an essential element for mice to initiate a protective immunity against Vaccinia virus. FEMS Microbiol Ecol 2016;92:1–8. https://doi.org/10.1093/femsec/fiv147.

[33] Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al. Microbiota regulates immune defense against respiratory tract influenza A virus infection. Proc Natl Acad Sci U S A 2011;108:5354–9. https://doi.org/10.1073/pnas.1019378108.

[34] Abt MC, Osborne LC, Monticelli LA, Doering TA, Alenghat T, Sonnenberg GF, et al. Commensal bacteria calibrate the activation threshold of innate antiviral immunity. Immunity 2012;37:158–70. https://doi.org/10.1016/j.immuni.2012.04.011.

[35] Wang J, Li F, Sun R, Gao X, Wei H, Li L-J, et al. Bacterial colonization dampens influenza-mediated acute lung injury via induction of M2 alveolar macrophages. Nat Commun 2013;4:2106. https://doi.org/10.1038/ncomms3106.

[36] Wang J, Li F, Wei H, Lian Z-X, Sun R, Tian Z. Respiratory influenza virus infection induces intestinal immune injury via microbiota-mediated Th17 cell–dependent inflammation. J Exp Med 2014;211:2397–410. https://doi.org/10.1084/jem.20140625.

[37] Deriu E, Boxx GM, He X, Pan C, Benavidez SD, Cen L, et al. Virus affects intestinal microbiota and secondary salmonella infection in the gut through Type I interferons. PLoS Pathog 2016;12.. https://doi.org/10.1371/journal.ppat.1005572.

[38] Andrade AC do SP, Lima MT, Oliveira GP, Calixto RS, de Sales e Souza L, de Souza D da G, et al. Daily ingestion of the probiotic Lactobacillus paracasei ST11 decreases Vaccinia virus dissemination and lethality in a mouse model. Benef Microbes 2017. https://doi.org/10.3920/BM2016.0074.

[39] Wang X, Hu W, Zhu L, Yang Q. *Bacillus subtilis* and surfactin inhibit the transmissible gastroenteritis virus from entering the intestinal epithelial cells. Biosci Rep 2017. https://doi.org/10.1042/BSR20170082.

[40] Hummelen R, Hemsworth J, Changalucha J, Butamanya NL, Hekmat S, Habbema JDF, et al. Effect of micronutrient and probiotic fortified yogurt on immune-function of anti-retroviral therapy naive HIV patients. Nutrients 2011. https://doi.org/10.3390/nu3100897.

[41] Nagata S, Asahara T, Ohta T, Yamada T, Kondo S, Bian L, et al. Effect of the continuous intake of probiotic-fermented milk containing Lactobacillus casei strain Shirota on fever in a mass outbreak of norovirus gastroenteritis and the faecal microflora in a health service facility for the aged. Br J Nutr 2011. https://doi.org/10.1017/S000711451100064X.

[42] Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health. Front Physiol 2015;6.. https://doi.org/10.3389/fphys.2015.00081.

[43] Klebanoff SJ, Coombs RW. Viricidal effect of Lactobacillus acidophilus on human immunodeficiency virus type 1: possible role in heterosexual transmission. J Exp Med 1991;174:289–92. https://doi.org/10.1084/jem.174.1.289.

[44] Conti C, Malacrino C, Mastromarino P. Inhibition of herpes simplex virus type 2 by vaginal lactobacilli. J Physiol Pharmacol 2009;60:19–26.

[45] Yildirim S, Yeoman CJ, Janga SC, Thomas SM, Ho M, Leigh SR, et al. Primate vaginal microbiomes exhibit species specificity without universal Lactobacillus dominance. ISME J 2014;8:2431–44. https://doi.org/10.1038/ismej.2014.90.

[46] Ongradi J, Ceccherini-Nelli L, Pistello M, Specter S, Bendinelli M. Acid Sensitivity of Cell-Free and Cell-Associated HIV-1: Clinical Implications. AIDS Res Hum Retroviruses 1990;6:1433–6. https://doi.org/10.1089/aid.1990.6.1433.

[47] Kempf C, Jentsch P. Inactivation_of_Human_Immunodeficiency_Virus_HIV_.17 (2).pdf 1991:828–30.

[48] Martin LS, McDougal JS, Loskoski SL. Disinfection and Inactivation of the Human T Lymphotropic Virus Type III/Lymphadenopathy-Associated. Virus 1985:400–4003. https://doi.org/10.1093/infdis/152.2.400.

[49] O'Connor TJ, Kinchington D, Kangro HO, Jeffries DJ. The Activity of Candidate Virucidal Agents, Low pH and Genital Secretions against HIV-1 In Vitro. Int J STD AIDS 1995;6:267–72. https://doi.org/10.1177/095646249500600409.

[50] Connor RI. Sensitivity of non-clade B primary HIV-1 isolates to mildly acidic pH. J Acquir Immune Defic Syndr 2006;43:499–501. https://doi.org/10.1097/01.qai.0000243048.50451.18.

[51] Tuyama ACG, Cheshenko N, Carlucci MJ, Li J-H, Goldberg CL, Waller DP, et al. ACIDFORM inactivates herpes simplex virus and prevents genital herpes in a mouse model: optimal candidate for microbicide combinations. J Infect Dis 2006;194:795–803. https://doi.org/10.1086/506948.

[52] Mitra A, MacIntyre DA, Marchesi JR, Lee YS, Bennett PR, Kyrgiou M. The vaginal microbiota, human papillomavirus infection and cervical intraepithelial neoplasia:

what do we know and where are we going next? Microbiome 2016;4:58. https://doi.org/10.1186/s40168-016-0203-0.
[53] Straight SW, Herman B, McCance DJ. The E5 oncoprotein of human papillomavirus type 16 inhibits the acidification of endosomes in human keratinocytes. J Virol 1995;69:3185–92.
[54] Brotman RM, Shardell MD, Gajer P, Tracy JK, Zenilman JM, Ravel J, et al. Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis 2014. https://doi.org/10.1093/infdis/jiu330.
[55] Miller EA, Beasley DAE, Dunn RR, Archie EA. Lactobacilli dominance and vaginal pH: Why is the human vaginal microbiome unique? Front Microbiol 2016;7.. https://doi.org/10.3389/fmicb.2016.01936.
[56] Cohen CR, Lingappa JR, Baeten JM, Ngayo MO, Spiegel CA, Hong T, et al. Bacterial vaginosis associated with increased risk of female-to-male HIV-1 transmission: a prospective cohort analysis among african couples. PLoS Med 2012;9:18. https://doi.org/10.1371/journal.pmed.1001251.
[57] Cherpes TL, Meyn LA, Krohn MA, Lurie JG, Hillier SL. Association between acquisition of herpes simplex virus type 2 in women and bacterial vaginosis. Clin Infect Dis 2003;37:319–25. https://doi.org/10.1086/375819.
[58] McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, Jaoko W, et al. Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. Lancet Infect Dis 2018. https://doi.org/10.1016/S1473-3099(18)30058-6.
[59] Passmore JAS, Jaspan HB, Masson L. Genital inflammation, immune activation and risk of sexual HIV acquisition. Curr Opin HIV AIDS 2016. https://doi.org/10.1097/COH.0000000000000232.
[60] Cohn JA, Hashemi FB, Camarca M, Kong F, Xu J, Beckner SK, et al. HIV-inducing factor in cervicovaginal secretions is associated with bacterial vaginosis in HIV-1-infected women. J Acquir Immune Defic Syndr 2005. https://doi.org/10.1097/01.qai.0000146599.47925.e0.
[61] Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, et al. Bacterial communities in women with bacterial vaginosis: High resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS ONE 2012. https://doi.org/10.1371/journal.pone.0037818.
[62] Clarke MA, Rodriguez AC, Gage JC, Herrero R, Hildesheim A, Wacholder S, et al. A large, population-based study of age-related associations between vaginal pH and human papillomavirus infection. BMC Infect Dis 2012;12:33. https://doi.org/10.1186/1471-2334-12-33.
[63] Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 2011;1. https://doi.org/10.1101/cshperspect.a006841.
[64] Hladik F, McElrath MJ. Setting the stage: host invasion by HIV. Nat Rev Immunol 2008;8:447–57. https://doi.org/10.1038/nri2302.
[65] Thrall PH, Antonovics J, Bever JD. Sexual transmission of disease and host mating systems: within-season reproductive success. Am Nat 1997;149:485–506. https://doi.org/10.1086/521238.
[66] Thrall PH, Antonovics J, Dobson AP. Sexually transmitted diseases in polygynous mating systems: prevalence and impact on reproductive success. Proc R Soc B Biol Sci 2000;267:1555–63. https://doi.org/10.1098/rspb.2000.1178.
[67] Lockhart AB, Thrall PH, Antonovics J. Sexually transmitted diseases in animals: ecological and evolutionary implications. Biol Rev Camb Philos Soc 1996;71:415–71. https://doi.org/10.1111/j.1469-185X.1996.tb01281.x.

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].

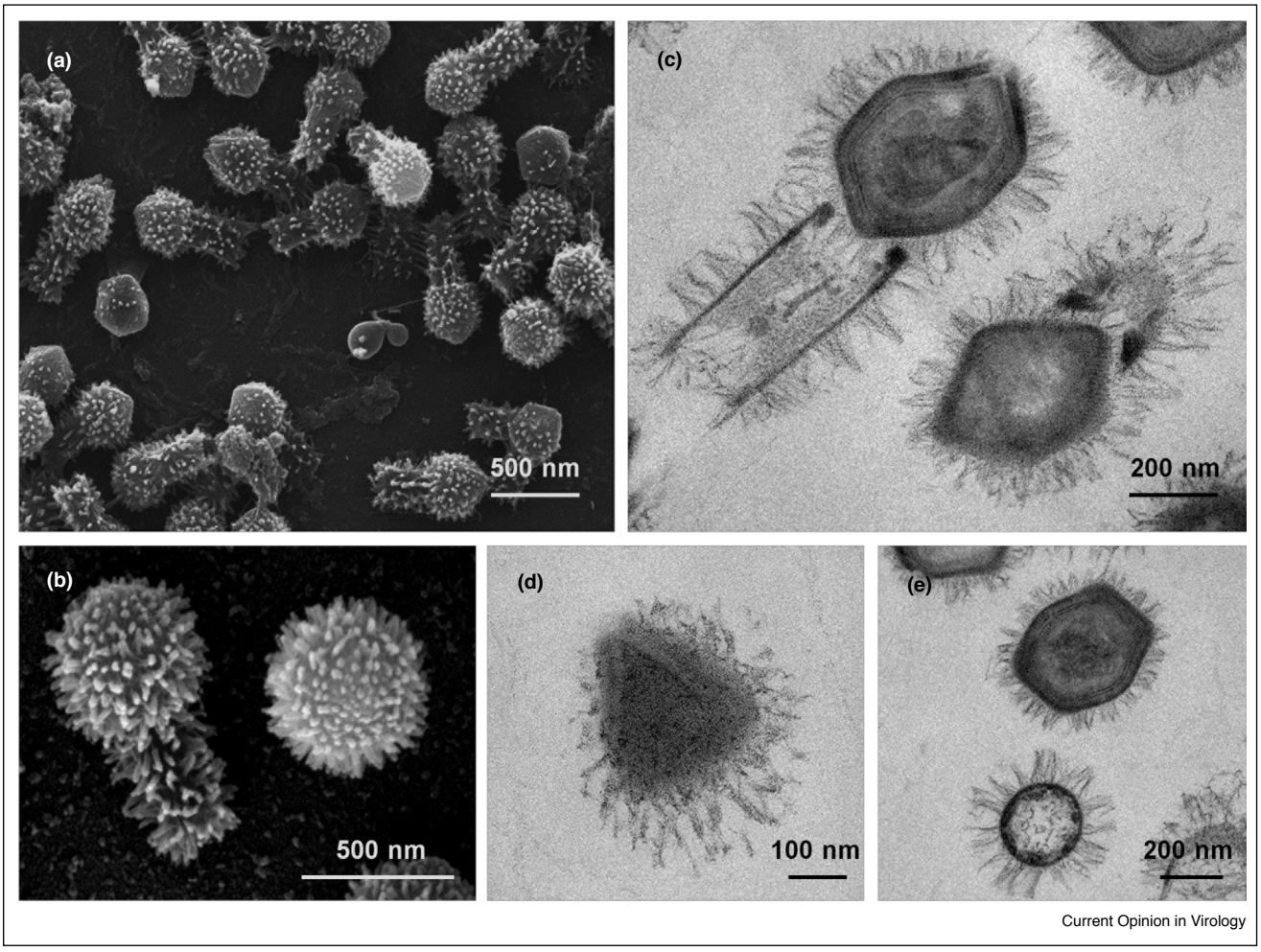
Even more striking is the structure observed for tupanviruses, a new group of viruses within the family *Mimiviridae*. These viruses were recently isolated from extreme environments in Brazil and are capable of infecting a wide variety of amoebae species [22••]. Tupanvirus has more than 1200 genes, and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

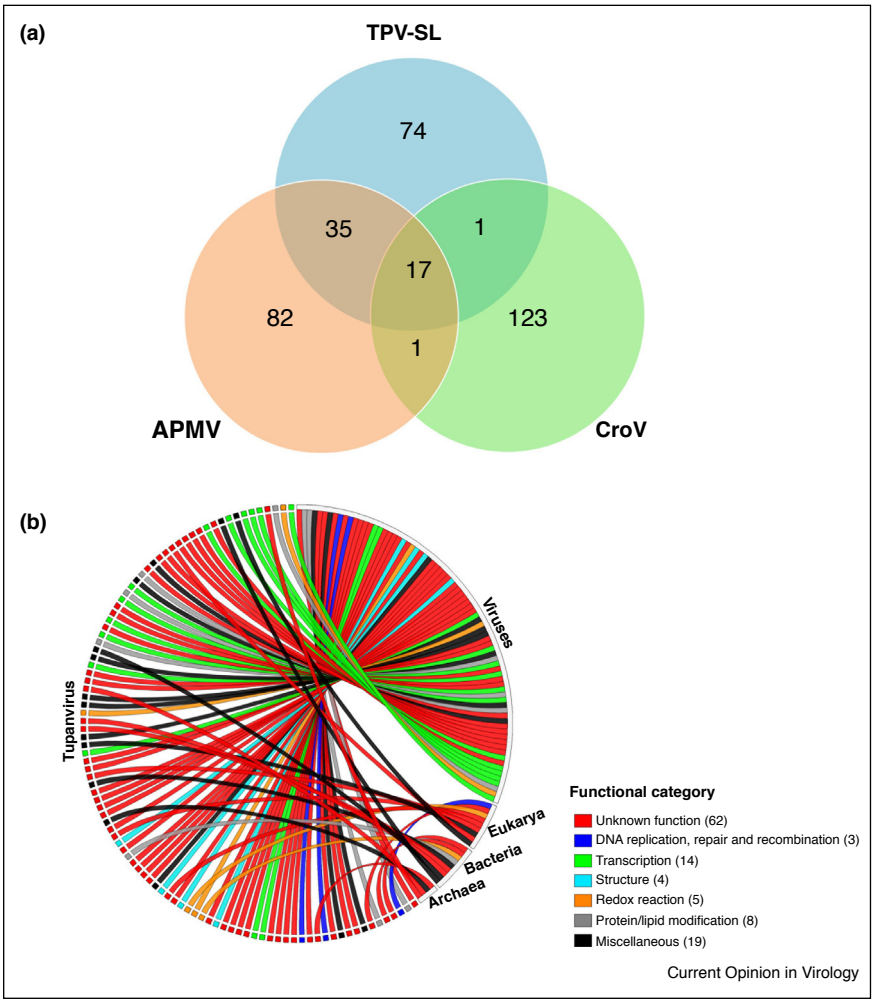
According to many proteomic studies within the known virosphere, the mature particles of amoebal giant viruses are composed of a great number of proteins [4,24••,25,26•]. This has generated hypotheses that seek to explain why only a small fraction of the nearly one thousand proteins encoded by mimiviruses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



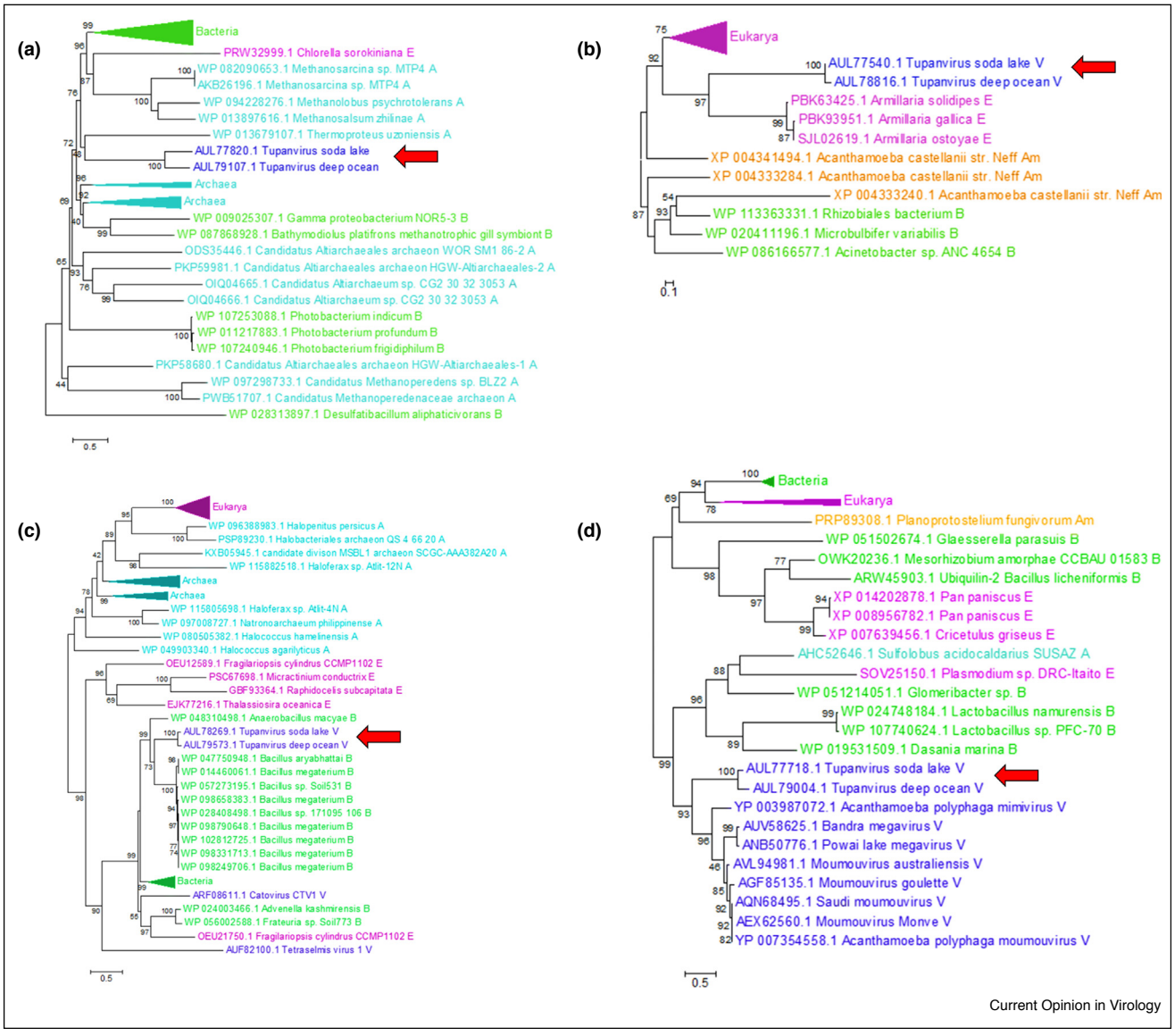
Proteome analysis of tupanvirus soda lake.
(a) Venn diagram of a comparative analysis of the viral particle proteome of TPV-SL, APMV, and CroV containing 127, 136, and 141 proteins, respectively. The analysis was performed using the proteomic data of each virus obtained from the literature [21,25,31]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

Title: Giant viruses and their parasites

Authors information:

Author: Rodrigo Araújo Lima Rodrigues

rodriguesral07@gmail.com

Co-author: Ana Cláudia dos Santos Pereira Andrade

anaclaudiaandrade29@gmail.com

Co-author: Graziele Pereira Oliveira

graziufmg2008@gmail.com

Corresponding author: Jônatas Santos Abrahão

jonatas.abrahao@gmail.com

Affiliation: Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil – postal code 31270-901.

Abstract

Giant viruses were identified from the beginning of the 21st century and broke many paradigms of traditional virology related to particle size and genome content. These viruses are associated with different amoebae species

and have been found in various places on the planet. Due to their astonishing complexity, giant viruses lie in the center of a hot debate about the origin and evolution of viruses. Moreover, they can be infected by other viruses, named virophages, and have their own genetic mobilome. In this chapter, we describe the history and general features of the different giant viruses and their parasites.

Keywords

Amoebae, evolution, genomics, giant virus, megavirales, mimivirus, mobilome, NCLDVs, proteomics, replication cycle, virophage.

Glossary

Megavirales: A proposed taxon (order) to comprise all nucleo-cytoplasmic large DNA viruses, given its hypothetical common origin.

MIMIVIRE: A mimivirus defense system against virophage infection.

Polintovirus: Large DNA transposons found in cellular organisms that encodes up to 10 proteins, including a homolog of viral capsid, possibly generating infectious particles.

Provirophage: A virophage genome integrated in the host DNA.

Transpoviron: Named after the “transposon,” this is a small linear DNA transposable element found in mimivirus genomes.

Viral factory: A viral-induced region in the host cytoplasm, wherein genome replication and viral morphogenesis of giant viruses occur.

Virophage: A new class of viruses associated with giant viruses, depending on the viral host for replication.

Introduction

Viruses are traditionally conceived as filterable agents capable of passing through membranes with pore sizes of 0.22 µm, containing small genomes encoding only a few proteins. However, with the discovery of giant viruses, these concepts have been broken down, and a new era of virology has arisen. Giant viruses are part of a group of nucleo-cytoplasmic large DNA viruses (NCLDV), in which the viral particle size and structure and the genome length and complexity make them special within the group. In the original NCLDV proposal, the group consisted of the viral families *Poxviridae, Asfarviridae, Iridoviridae* and *Phycodnaviridae*. These families are composed of viruses that contain large double-stranded DNA (dsDNA) genomes, which replicate partially or entirely into the host cytoplasm, and which also share genetic similarities that points to a hypothetical common origin, thus comprising a monophyletic group. However, NCLDV only became a more intense topic of discussion and evolutionary study in 2003, due to the discovery of the first giant virus of

amoebae named Acanthamoeba polyphaga mimivirus (APMV). Phylogenetic analyses indicated that APMV could be categorized with other viruses belonging to the NCLDV group, and the fact that this virus occupied an isolated branch in the phylogenetic tree that was associated with distinct morphological features led to the establishment of a new family, named *Mimiviridae*. The discovery of mimiviruses paved the way for the isolation of other giant viruses, with the development of new isolation techniques that were no longer constrained by the classical definitions of viruses based on particle size. As a consequence, new NCLDV members were discovered, among which included marseilleviruses, faustoviruses, pandoraviruses, cedratviruses, pithoviruses, orpheovirus, mollivirus, pacmanvirus and kaumoebavirus, and new definitions of viruses are currently being debated. These viruses exhibit many distinctive features, along with some common characteristics that classify them as part of the NCLDV group, which was proposed to comprise a new viral order named "Megavirales." In this chapter we describe how the giant viruses were discovered and compiled the basic features about them, regarding diversity, replication cycle, genomics and evolution. Finally, we talk about the genetic mobilome and virophages, the parasites of giant viruses.

The serendipitous discovery of amoebae giant viruses

The isolation of the first mimivirus accidentally occurred during studies of amoebae-associated pathogens from environmental water samples in the midst of a pneumonia outbreak in Bradford, England. This virus was initially considered to be a bacterium of amoebae, mainly due to its Gram-positive

staining back in 1992 that revealed a microorganism referred to as “Bradford coccus.” In addition, molecular attempts had failed to identify the new isolate. It was years later, when the sample was taken to Aix Marseille University where it was characterized and described in 2003 as the largest virus defined at that time. Its classification as a virus was confirmed by the observation of the eclipse-phase during its replication and the visualization of a likely icosahedral-shaped particle via transmission electron microscopy (TEM), which is typical of viruses, that was also covered in long fibrils. The virus was named after its ability to mimic microbes and the host cell it was first isolated from, *Acanthamoeba polyphaga*. Mimivirus particles measure about 750 nm and are composed of a peculiar structure, characterized by a protein capsid containing an internal lipid membrane and surrounded by a dense glycoprotein fibril layer that is important for viral attachment to the host cell and other microorganisms (Fig. 1A). The capsid does not exhibit the typical icosahedral symmetry, presenting instead a modified vertex in starfish shape, named *stargate*, a portal from where the viral genome is released into the hosts’ cytoplasm (Fig. 1B). Its genome consists of a long and complex linear double-stranded DNA molecule of approximately 1.2 Mbp, encoding over 1,000 genes with many different functions, such as enzymes involved in DNA replication, recombination and repair (e.g., DNA polymerase and DNA topoisomerase), transcription (RNA polymerase and mRNA capping enzyme) and translation (translation factors and aminoacyl-tRNA synthetases), among others. The mimiviruses penetrate into the host cell (free-living amoebae) by phagocytosis; in principal, this generates rounding due to a cytopathic effect, followed by lysis of the amoebae. Similar to other viruses belonging to the NCLDV group, mimiviruses replicate in

the cytoplasm of the host cell, where they form large viral factories.

<Figure 1 near here>

Since the discovery of APMV in the context of pneumonia, discussions about the possible relationship between mimiviruses and this disease have been raised. One study showed that laboratory mice inoculated intracardially with mimivirus displayed histopathologic symptoms, and mimiviruses were re-isolated from lung samples of these animals. Moreover, patients with pneumonia presented antibodies against mimivirus, and a high positivity of mimivirus in hospital facilities was recorded. Years later, the isolation of mimivirus from stool and bronchoalveolar lavage samples from patients with pneumonia strengthened the relationship between mimiviruses and this disease. However, human cells capable of being infected by these viruses, wherein the virus could establish a productive replication cycle, have not been found yet. Recent studies found no evidence of mimivirus replication in human peripheral blood mononuclear cells, despite its ability to enter and interfere with antiviral responses. Furthermore, other studies revealed the absence of mimiviruses in patients with respiratory disease by means of molecular biology and/or serology. However, it is important to consider the possibility that mimivirus might be a member of the vertebrates' virome, since their genome has been detected in humans, as well as both in domesticated and wild mammals. Nonetheless, this association needs more evidence to be confirmed. It is worth noting that a mimivirus-related virus infecting sturgeons has been described in the last years, causing lethal disease in the integumentary systems

of these animals. Further studies are required to better evaluate the putative pathogenic role of mimiviruses and other giant viruses in humans and other vertebrates.

Diversity and distribution of giant viruses in the world

After the discovery of mimiviruses, other amoebae giant viruses have been described over the past decade in different places around the world, revealing an increasingly complex fraction of the virosphere. These viruses have a great structural diversity, with particles of different sizes and shapes, ranging from icosahedral capsids of about 200 nm to circular, ovoid or tail-containing particles that exceed 1,000 nm in length. Four distinct groups, named marseilleviruses (family *Marseilleviridae*), kaumoebavirus, pacmanvirus and faustovirus (all phylogenetically related to *Asfarviridae* members), are composed of smaller viruses that present icosahedral capsids of 175–250 nm. Marseilleviruses were the second group of amoeba viruses to be discovered and were associated with *Acanthamoeba* cells in water samples from Paris, France. Over the years, other isolates were described, expanding the new viral family that is currently composed of five different lineages (A–E), with viruses discovered in different regions of the planet, such as Brazil, Tunisia and Japan, among others. The other groups, despite infecting amoeba cells, are most closely related to the African swine fever virus, a mammalian parasite. Pacmanvirus also infects *Acanthamoeba* cells, but kaumoebavirus and faustoviruses are associated with *Vermamoeba vermiformis* cells, a different amoebal host. Similar to mimiviruses, these smaller viruses have an internal

lipid membrane surrounding their genetic material. Electron microscopy images demonstrate an overall similar structure between these different viruses (Fig. 2). Although in-depth structural analysis of giant viruses particles are scarce, some studies have indicated exclusivities for some viruses. Among them, faustovirus has the peculiarity of being formed by a double concentric capsids that connect to each other by protrusions of the internal capsid surface, which was not observed for other viruses until now.

Other groups have larger particles that are ovoid in shape. In 2013, the description of pandoraviruses set a new limit to the viral particles' sizes, with ovoid-shaped virions with 1.0 µm presenting an ostiole-like apex at one end of its capsid (Fig. 2). These viruses were first isolated in Chile and Australia, but novel isolates have recently been found in Germany, France, New Caledonia and Brazil, constituting an expanding putative Pandoraviridae family. Curiously, the pandoravirus capsid composition remains unknown, since no protein even remotely similar to the major capsid protein (MCP) was detected in the genomes of these viruses. A year later, the previous size limit was pushed once again by the discovery of the giant Pithovirus sibericum, isolated from 30,000-year-old permafrost soil samples from Siberia, composed of ovoid particles reaching more than 1.5 µm in length, making them the largest viral particles known so far. The particles have a striated capsid wall and a hexagonal grid-like structure, called a cork, which is generally found at one extremity of the particle and is an analogous structure to the pandoraviruses' apex. Curiously, a new virus was isolated from the same ancient sample, named Mollivirus sibericum, which is a spherical-shaped virus that is approximately 600 nm in diameter, but no further studies have been performed to date. A contemporary pithovirus was

isolated in France that exhibited a similar structure and a genomic conservation, named Pithovirus massiliensis. Due to these structurally and genomically distinct features, a new viral family called "*Pithoviridae*" was suggested. In 2016, cedratvirus, a virus with a morphology similar to the pithovirus, was described, but it contained smaller particles with a mean size of 1.2 µm. Like the pithovirus, cedratvirus has apical corks, but at both ends of the particles. Other cedratviruses have been isolated in France and Brazil, with the Brazilian isolate exhibiting differences in particle and genome size and probably constituting a new lineage among cedratviruses (Fig. 2). Cedratviruses and pithoviruses are phylogenetically close and might constitute a new single family. A distantly related *Pithoviridae* member named Orpheovirus was recently isolated using *V. vermamoeba* cells, which is morphologically similar to pandoraviruses, however, have more electron-dense capsid layers that are covered with short fibrils on their outer surface (Fig. 2). More studies should be performed to get a clear picture of the diversity and distribution of these new giant viruses.

<Figure 2 near here>

Currently, the family *Mimiviridae* encompasses the largest number of representatives described to date, which have considerable differences in both structure and genomic features. After the discovery of APMV, several other mimivirus-like viruses were isolated from different parts of the world that exhibited structural similarities, but with genomic differences, contributing to the formation of three different lineages, called lineage A (represented by APMV), B (represented by Moumouvirus) and C (represented by Megavirus). Distantly

related mimiviruses were also described infecting marine flagellates, named Cafeteria roenbergensis virus (CroV), a new virus exhibiting large particles with many genes shared with mimiviruses, which put it on a new viral genus, the *Cafeteriavirus*, within the family *Mimiviridae*. The bodo saltans virus, a kinetoplastid-infecting virus, constitutes a new member of the family, the first isolated member of the putative subfamily "*Klosneuvirinae*", whose members were initially identified by metagenomics analysis, revealing a new group of mimiviruses with unexpected genomic features, such as a much larger translation-related gene set. More recently, the structural and genomic complexity of the family was expanded with the discovery of Tupanviruses in samples collected from extreme places in Brazil, which are the viruses with the most complete translational apparatus in the virosphere. The viral particles are formed by a capsid similar to that of APMV and are associated with a cylindrical tail of about 550 nm in length and 450 nm in diameter, with the whole particle covered in fibrils (Fig. 2). The size of these particles is approximately 1.2 μm on average, although some particles may be up to 2.3 μm, due to the plasticity of the tail.

The discovery of the first viruses associated with amoebas at the beginning of 21st century, with their diverse and complex structural characteristics, led to an intense search for new giant viruses in distinct places of the world. Since then, the number of new isolates has grown every year. These viruses have already been isolated using samples collected on each continents from various conditions, such as polluted environments, those with extreme conditions of temperature and pH, slightly anthropized regions and clinical samples. The constant discovery of giant viruses in wide-ranging types

of samples reinforces the idea that these viruses are ubiquitous, as well as their protist hosts. Metagenomic studies reinforce the ubiquitousness, further indicating that many amoeba-related viruses have yet to be discovered. Remarkably, a large metagenomic analysis of soil samples from the Havard forest (Massachusetts, USA) identified the complete genome of 16 new giant viruses, representing new lineages of klosneuviruses, other members of the family *Mimiviridae* and hundreds of MCP fragments belonging to giant viruses, thus indicating a huge diversity of these viruses in soil ecosystems that remain to be characterized. The host spectrum for most of these viruses is restricted to a few species of amoebae of the genus *Acanthamoeba* and/or *Vermoameba*. However, recent discoveries of new members of the *Mimiviridae* family have shown that the spectrum may be broader, since they can infect different hosts, like other groups of protozoa and fish species (e.g., Bodo saltans virus, tupanviruses and namao virus). Novel studies involving prospection in unexplored regions using different cell platforms and isolation techniques could reveal exciting new viral groups, contributing to an increase in our knowledge about the viruses' diversity, ecology and evolution.

Replication cycle: Inside the life style of the giants

The replication cycle of giant viruses begins with viral entry into the host cells. For most of these viruses, this process has been proposed to occur by phagocytosis. Drugs such as cytochalasin, a phagocytosis inhibitor, have already been used to demonstrate the impact of this entry mechanism for mimivirus, cedratvirus and large membranous vesicles containing marseillevirus

particles. For pandoravirus, pithovirus, mollivirus, orpheovirus, faustovirus, pacmanvirus and kaumoebavirus, entry by phagocytosis was suggested by electron microscopy image analyses. The phagocytosis process requires the recognition of particles of at least 500 nm, so the assumption that giant viruses enter by this mechanism seems reasonable. However, other smaller viruses, such as marseilleviruses (~ 250 nm) do not fulfill this criteria; thus, other mechanism are likely involved. A thorough investigation of the replication cycle of this virus demonstrated that it is able to penetrate in cells by two alternative routes: by phagocytosis of grouped particles inside giant vesicles, or by exploiting other endocytic pathway for single particles. It is not clear yet if a similar process could occur for other large (yet not giant) viruses, such as faustoviruses and kaumoebavirus. After entry, the viral genome is released into the host cytoplasm. It has been observed that the uncoating process of mimivirus, mollivirus, pandoravirus, pithovirus, faustovirus and cedratvirus occurs after the fusion of the inner viral membrane with the endosomal membrane, allowing the formation of a channel by which the genome is released into the host cell. For other giant viruses, the uncoating process has not been described in detail yet. Mimiviruses release their genomes after stargate opening, and it was verified that endosome acidification is important for the induction of this step. The genomes of pandoraviruses, orpheovirus, pithovirus and cedratvirus are released by their respective ostiole-like apex and corks at the extremities of the particles.

Following the release of the genome, a typical viral eclipse phase is observed, wherein no viral particles are detected in the host cytoplasm, followed by the formation of small viral factories (VFs) in the cytoplasm. Throughout the

infection, these VFs increase in size and occupy a large portion of the cellular cytoplasm in the later stages of infection. VFs of most giant viruses are non-delimited electron-lucent areas that contain particles at different stages of morphogenesis. Nevertheless, mimivirus VFs are well-defined structures, with at least two areas described: the central area, where replication of the genome and assembly of the structures necessary to form the viral capsid occurs, and a less electron-dense region of apparent fibrillar nature in the periphery. This outermost portion was named the fibrils acquisition area, since it was observed that is in this region the newly formed particles acquire the fibrils. Such compartmentalization of VFs was not observed for other giant viruses so far. Previous studies proposed that capsid morphogenesis of mimiviruses, pandoraviruses and cedratvirus is initiated by crescent-shaped lamellar structures that can be observed inside the VFs. The morphogenesis of cedratvirus involves crescent-shaped structures that assume staple-shaped and horseshoe conformations, depending on the section of TEM. Initially, only one of the corks is observed; then the capsid is filled and closed with the emergence of the second cork, and the capsid wall becomes thicker. Regarding pandoravirus, the capsid and internal contents seem to be assembled simultaneously. A pattern of where the assembly of the particle begins was not observed, since it can be at the extremity that presents ostiole-like apex or at the opposite end. For mimiviruses, capsid precursors increase in complexity and assume a pseudo-icosahedral symmetry. Then genome acquisition occurs at the stargate on the opposite side, and this step may occur simultaneously with the acquisition of surface fibrils.

During the morphogenesis step, defective particles have been recurrently observed for different groups of viruses, including mimiviruses, pandoraviruses, cedratviruses, among others. For mimiviruses, these unusual particles were initially associated with the presence of virophages, parasites of giant viruses (described later in this chapter). However, recent analysis demonstrated the formation of defective particles, even in the absence of these parasites, thus, making clear that malformed particles are a natural process during the replicative cycle of these viruses. Once viral particles are fully assembled, release of most of these viruses occurs after cell lysis. Marseillevirus can release many particles into giant vesicles after cell lysis, which boosts its entry into another cell host by exploiting different entry strategies. Moreover, it has been suggested that the exocytosis process could be used for releasing mollivirus, pandoravirus, orpheovirus and cedratvirus particles. However, the viral release through exocytosis still needs to be further investigated for the majority of these viruses. The real impact of this viral release mechanism is still not well known.

Overall, in spite of some peculiarities of each viral group, a general picture of the replication cycle of giant viruses is observed and composed by different steps (Fig. 3). Viruses enter host cells through endocytosis or phagocytosis (i), release the genome into the cytoplasm (ii), establish a viral factory (iii) where the genome is replicated (iv), new particles are assembled during the morphogenesis step (v), and finally, the new progeny is release by cell lysis or exocytosis (vi). It is still not clear if these viruses depend on the host nucleus to complete their replication cycle, but some evidence for this has been described for marseilleviruses, pandoraviruses, mimivirus and mollivirus.

Additional analysis should be done to better elucidate this important biological feature for the different groups of giant viruses.

<Figure 3 near here>

Genomics and proteomics of giant viruses

A striking feature of giant viruses is the presence of an extensive dsDNA genome containing hundreds of genes. Some viruses have a linear genome, such as mimiviruses, tupanviruses, pandoraviruses, mollivirus and pacmanvirus, while others have a circular genome, such as marseilleviruses, fautoviruses, kaumoebavirus, pithoviruses, cedratviruses and orpheovirus. For the most part, the genome of these viruses is A/T-rich, ranging from ~ 55% in marseillevirus to ~ 75% in mimiviruses. Pandoravirus and mollivirus differ from the others in this regard, presenting genomes with G/C content above 60%. The size of the genome varies among the different viral groups between, ~ 350 kb to ~ 2500 kb, with Pandoravirus salinus containing the largest viral genome described so far. These viruses present, on average, 1 gene/Kb and short intergenic regions of approximately 200 nucleotides, although large regions of thousands of nucleotides can occur between genes.

Even more impressive than the size of the genome is the genetic content of these viruses. Sequencing of the APMV genome in 2004 revealed the presence of genes rarely found in the virosphere, some of which are considered exclusive to the cellular world. Giant viruses present several genes involved in DNA replication, recombination and repair, nucleotide and carbohydrate

374 metabolism, transcription (including RNA polymerases and various transcription
375 factors) and protein synthesis. In the latter category, members of the family
376 *Mimiviridae* are characterized by exhibiting several transfer RNAs (tRNAs),
377 many translation factors (involved in the three stages of the process, i.e.,
378 initiation, elongation and release) and aminoacyl-tRNA synthetases (aaRS),
379 which are key enzymes in the process of protein synthesis and have never
380 been observed before for other viruses. In this context, the tupanviruses are the
381 viruses that present the largest arsenal of genes involved in the translation
382 processes observed in the virosphere, with up to 70 tRNAs and 20 aaRS (one
383 for each proteinogenic amino acid encoded by the standard genetic code).
384 Other giant viruses also have components of the translational apparatus, but in
385 much smaller quantity and diversity. Although several translation-related genes
386 exist in these viruses, no ribosomal RNA (rRNA) gene has been detected in a
387 virus thus far, making them still dependent on their hosts to synthesize proteins.
388 In spite of several gene novelties observed, a common feature in all giant
389 viruses at the time of their discovery is the large number of genes with unknown
390 functions, without any similarity to other genes deposited in databases, called
391 ORFans, with values varying from 31 to 84% of ORFans in the genome.
392 Pandoravirus salinus has the larger fraction of ORFans among the giant viruses
393 (Fig. 4). This massive number of new genes reflects how different these viruses
394 are and suggests that they are a valuable sources of genetic diversity.
395
396 <Figure 4 near here>
397

Little is known about the transcriptional profile of these viruses and how the regulation of gene transcription occurs. Conserved A/T-rich sequences were identified in most of the mimivirus, marseillevirus, faustovirus and kaumoebavirus genes. Analyses based on RNA sequencing for APMV indicated the existence of a temporal profile of gene expression, with the AAAATTGA promoter sequence being associated with early expressed genes. Other members of the family *Mimiviridae* exhibit the same promoter motif, such as CroV, Bodo saltans virus and tupanvirus, which suggests it being a typical feature of this viral family. Marseillevirus also has genes expressed at different times in the replication cycle, but the previously identified motif promotor (AAATATTT) does not appear to be specifically related to any temporal class of genes, as is observed for mimiviruses. Most genes involved in cell signaling (e.g., serine/threonine kinases) and DNA replication and repair (e.g., DNA polymerase and DNA topoisomerases) are expressed at earlier times in the viral replication cycle, whereas transcription, translation and especially those involved in viral morphogenesis, such as constituents of the capsid (e.g., MCP and DNA packaging ATPase), are expressed later during the infectious cycle. In addition, transcripts of these viruses are polyadenylated, and hairpin polyadenylation signals have been identified throughout the genome of the mimivirus; however, there is still no information for the other giant viruses, being a broad field for new investigations.

Proteomics studies have revealed that giant virus particles are composed of a large amount of proteins, with more than 100 proteins found in the particles of APMV, CroV, Tupanvirus soda lake, Pithovirus sibericum, Mollivirus sibericum, pandoraviruses and E12 faustovirus. In the particles of all these

viruses proteins with different functions were detected that were involved in DNA replication, oxidative pathways, lipid modification, transcription processes, in addition to the already expected structural genes. It is worth noting that in all viral particles, products from ORFans were also identified, demonstrating that new genes discovered in these entities are truly expressed and translated, although their exact function remains unknown. Many genes identified in viral particles, as well as others expressed during the viral replication cycle, have multiple origins (i.e., from viruses, eukaryotes, bacteria and archaea), making giant DNA virus true genetic mosaics. This feature is observed in several groups of giant viruses, but is especially evident in the marseilleviruses, where a large fraction of their genes originate from lateral gene transfer events. This genomic mosaicism is attributed in large part to the sympatric lifestyles of these organisms (i.e., from the same environment as several other microorganisms), since they infect amoebas, and they are considered melting pots for the emergence of new organisms.

Origin and evolution: an intriguing enigma

Giant amoeba viruses share several genes with other members of the NCLDV group, which is a possibly monophyletic group within the virosphere. With the discovery of different viral groups, the amount of so-called core genes has been drastically reduced, with only three genes currently present in members of all viral groups, named the D5 primase-helicase, viral late transcription factor 3 (VLTF3) and DNA polymerase B family. Some core genes appear to have been lost in some specific groups throughout evolution, such as

the MCP genes in pandoravirus and DNA packaging ATPase in pithovirus and cedratvirus. Phylogenetic analyses based on these core genes tend to maintain the same topology, with the presence of three major clades, with the different groups of amoeba viruses scattered among them (Fig. 5). Although they share a common ancestor with the other *Megavirales* members, it is possible that the gigantism observed for some groups has arisen independently throughout evolution, with the host acting as a selection agent of this characteristic. However, the origin and evolution of these viruses are at the heart of a heated debate that has lasted more than a decade, and there is still no scientific consensus.

<Figure 5 near here>

The discovery of the presence of an extensive genome (> 1.0 Mb) encoding hundreds of proteins (~ 1,000) in mimiviruses, with particular attention to translation genes like aaRS, contributed to reviving a long-standing debate about the nature of viruses (i.e., whether they are living organisms or not). Initial analyses suggested that mimiviruses could be remnants of a fourth domain of life, originating from a more complex organism (possibly a cell), which would have evolved through genomic reduction to adapt to the intracellular parasitism lifestyle. Additional phylogenomic analyses, associated with phylogenetic reconstructions based on protein fold conservation, reinforced this initial hypothesis, even inferring that giant viruses would have coexisted with primitive cellular life forms. However, alternative analyses involving conserved genes, as well as those related to the translation process, suggested an opposite

scenario, in which the giant viruses would have originated from simpler entities and which would expand the genome mainly through gene gain and duplication. These analyses used sequence sampling and varied methods of phylogenetic reconstruction and have indicated that the evolution of these viruses would have followed an accordion-like model, where there were losses and gains of genes along the process of evolution. A positive balance of this relation would have occurred, resulting in viruses with extensive genomes and giant particles. Thus, two diametrically opposed scenarios are in force and remain in constant debate among specialists.

The latter hypothesis has gained strength from comparative polintovirus analyses, a type of DNA transposon commonly identified in the genome of several cell organisms and which have genes homologous to some giant virus genes, including a major capsid protein and DNA polymerase B family genes. Recent analyses have suggested these mobile elements are the ancestors of NCLDVs, hence the various groups of giant amoeba viruses. These elements form a complex network of evolutionary interactions together with transpovirons and provirophages, elements that are part of the giant virus mobilome, something never before observed in other viruses.

Parasites of viruses: the nature of virophages and the genetic mobilome

A hallmark of giant viruses is the presence of their own genetic mobilome, which includes introns, transposable elements and even viruses, which are the so-called virophages. Virophages are a new group of viruses associated with giant viruses, the first isolate of which was described in 2008

498 and was associated with the Acanthamoeba castellanii mamavirus, a new
499 isolate of the family *Mimiviridae*. The new virus, called Sputnik virus, has
500 icosahedral particles of 75 nm and a circular A/T-rich genome of 18 Kb
501 encoding 21 ORFs (Fig. 6A, B). This virus was initially observed as associated
502 with mimivirus fibrils and was initially suggested as the mode of virophage entry
503 into amoebae cells. However, analyses based on other isolated virophages
504 have suggested that penetration may occur through clathrin-mediated
505 endocytosis. These viruses replicate in the viral factory of the giant virus,
506 appropriating from the viral host transcriptional machinery (Fig. 6A). These
507 viruses do not have their own RNA polymerase, and *in silico* analyses have
508 demonstrated the presence of promoter motifs similar to those found in
509 mimivirus, reinforcing the hypothesis that these viruses depend on the giant
510 virus host to have its genome transcribed, and subsequently, to form the new
511 progeny. The fact that the presence of the associated virus causes the
512 reduction of the multiplication of the giant virus by ~ 70%, coupled with other
513 peculiar characteristics of these parasites, differentiates them from the already
514 known satellite viruses leading to the establishment of a new category of virus.
515
516 <Figure 6 near here>
517
518 In the years following the discovery of the first isolate, other virophages
519 were isolated, such as the Mavirus that is associated with the Cafeteria
520 roebergensis virus, Sputnik 2, Sputnik 3, Zamilon and Rio Negro virus, and all
521 were associated with different isolates of mimiviruses. Interestingly, Sputnik 2
522 was identified integrated into the genome of a host mimivirus (Lentille virus),

which led to the creation of the term “provirophage.” Recently, the presence of provirophages in the genome of a unicellular algae (*Bigelowiella natans*), whose genes are actively transcribed, has been described, and their presence has led to the hypothesis that they would act as a defense system against giant virus infection. Mavirus was found integrated in multiple sites of the genome of the protozoan *Cafeteria roenbergensis*, and its activation was observed through CroV infection, which led to the generation of a progeny with subsequent cellular lysis. The newly produced virophages lead to the suppression of the giant virus replication and increase the survival of the cellular host community, reinforcing the hypothesis of an antiviral system triggered by the presence of a provirophage.

Interestingly, the Zamilon virophage is able to replicate with mimiviruses from lineage B and C, but not when in the presence of viruses from lineage A. Extensive analysis of the genome of different mimivirus isolates revealed that members from lineage A contain the insertion of a repeated Zamilon sequence within an operon, which was named the “mimivirus virophage resistant element” or MIMIVIRE. This system is composed of interspaced, repeated virophage sequences, along with nuclease genes, in a way analogous to the CRISPR-Cas system found in prokaryotes. By silencing the MIMIVIRE genes and the repeat sequences, the mimiviruses becomes susceptible to the virophage infection, which led to the proposition of a new nucleic-acid-based immunity system against virophage infection. An alternative explanation was further proposed, stating that the defense system is not CRISPR-like, but instead a protein interaction model would be responsible for the phenotype. Recently, a pivotal component in this system, the R354 gene of mimivirus encoding a nuclease,

548 had its structure defined, with analysis evidencing that the viral nuclease is
549 functionally similar to the Cas4 protein, thus providing additional evidence that
550 the MIMIVIRE is a new innate immune defense system.

551 Several other virophages have had their genomes identified by means of
552 metagenomic analyses in different places of the world, including Yellowstone
553 lake virophages (1–7, USA), Organic lake virophage and Ace lake mavirus
554 (Antarctic), and Dishui lake virophage and Qinghai lake virophage (China).
555 Other virophage genomes and gene fragments have been recovered from
556 metagenomics datasets from Lake Mendota and Trout Bog Lake (Wisconsin,
557 USA), identified as “freshwater virophage candidate genus.” Moreover,
558 virophages have also been identified in datasets from sheep rumen
559 metagenomics, constituting a hybrid group of virophages and polintons (linear
560 genomes with inverted terminal repeats and genes with a strong similarity to
561 polinton genes), named “sheep rumen virophages.” Some of these virophages
562 were associated with algae viruses, which would be responsible for controlling
563 the dynamics between the virus and algae-host in the environment. These new
564 viruses have a genome that ranges from 17–30 kb, encoding 16–34 ORFs and
565 containing several genes without known functions, but also containing some
566 conserved genes, such as major and minor capsid protein, protease and
567 packaging ATPase, which are used as genetic markers for these viruses and for
568 phylogenetic reconstructions (Fig. 6C). The presence of at least some of these
569 genes, along with the association or dependence on viruses of the NCLDV
570 group to fully replicate, characterize a virus as a virophage and are thus
571 classified in the family *Lavidaviridae*, which includes the genera *Sputnikvirus*
572 and *Mavirus* to date.

Virophages and provirophages are part of the genetic mobilome of giant viruses. Along with these elements are the transpovirons, a linear sequence of 7 Kb capable of integrating randomly into regions of the genomes of giant viruses, and which have been identified in different isolates of the three lineages of the genus *Mimivirus*. In addition, introns have already been detected in genes of different giant viruses, such as mimivirus and faustovirus, and together with inteins complete the known viral mobilome so far. Genomic analyses have indicated similarities between the mobile elements of giant viruses and polintons/mavericks, large transposable DNA elements found in various cellular organisms. Phylogenetic reconstructions suggest that virophages, as well as giant viruses, may have originated from polintoviruses, which constitute a complex network of evolutionary interaction.

Perspectives

Giant viruses and their parasites is a new and wide-spread field within virology. Since the discovery of the giant mimivirus, new viral groups have been described, amplifying the complexity of the virosphere as we know it, and making it necessary to advance the taxonomy of these viruses. New isolation techniques for subsequent viral characterization associated with metagenomic techniques have strongly contributed to the advancement of our understanding of this complex group of eukaryoviruses and their parasites. Finally, studies aimed at better understanding the biology of these viruses are necessary and may reveal new features that could be exploited from the biotechnological point

of view, given the enormous genomic and structural variety presented by these viruses, and the singular features of these giants of the virosphere.

Further reading

Abergel, C., Legendre, M., Claverie, J.M. (2015). The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiology Reviews* **39**, 779-96.

Abrahão, J., Silva, L., Silva, L.S., *et al*. (2018). Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nature Communications* **9**, 749.

Claverie, J.M. and Abergel, C. (2009). Mimivirus and its virophage. *Annual Review of Genetics* **43**, 49–66.

Colson, P., Levasseur, A., La Scola, B., *et al*. (2018). Ancestrality and Mosaicism of Giant Viruses Supporting the Definition of the Fourth TRUC of Microbes. *Frontiers in Microbiology* **9**, 2668.

Desnues, C., La Scola, B., Yutin, N., *et al*. (2012). Provirophages and transpovirons as the diverse mobilome of giant viruses. *Proceedings of the National Academy of Sciences of USA* **109**, 18078–18083.

Filée, J. (2013). Route of NCLDV evolution: The genomic accordion. *Current Opinion in Virology* **3**, 595–599.

Krupovic, M., Kuhn, J.H., Fischer, M.G. (2016). A classification system for virophages and satellite viruses. *Archives of Virology* **161**, 233–247.

La Scola, B., Desnues, C., Pagnier, I. *et al*. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* **455**, 100–4.

622 Raoult, D., Audic, S., Robert, C., *et al*. (2004). The 1.2-Megabase genome
623 sequence of Mimivirus. *Science* **306**, 1344–1350.
624 Xiao, C., Kuznetso, Y.G., Sun, S. *et al*. (2009). Structural studies of the giant
625 Mimivirus. *PLoS Biology* **7**, 0958–0966.
626 Yutin, N., Raoult, D., Koonin, E. V. (2013). Virophages , polintons , and
627 transpovirons : a complex evolutionary network of diverse selfish genetic
628 elements with different reproduction strategies. *Virology Journal* **10**, 1.
629 Yutin, N., Wolf, Y.I., Koonin, E. V. (2014). Origin of giant viruses from smaller
630 DNA viruses not from a fourth domain of cellular life. *Virology* **466–467**, 38–
631 52.

632

633 **Figure legends:**

634

635 **Figure 1) Electron microscopy images of mimivirus particles.** (A)
636 Transmission electron microscopy image demonstrating the capsid covered in
637 fibrils and an internal membrane; (B) scanning electron microscopy image
638 evidencing the stargate structure.

639

640 **Figure 2) Diversity of giant viruses.** Transmission and scanning electron
641 microscopy images of different viral groups, representing the large variety of
642 particle sizes and shapes among giant viruses. (A) mimivirus; (B) marseillevirus;
643 (C) pandoravirus; (D) tupanvirus; (E) orpheovirus; (F) cedratvirus. Scale bars:
644 100 nm.

645

Figure 3: Representative scheme of the amoeba giant viruses replication cycle. The particles can enter host cell by phagocytosis or endocytosis for individual particles of marseillevirus (I). Subsequently, the genome is released into the host cytoplasm (II). A viral factory is established in the cytoplasm of the (III) wherein the genome replicates (IV) and the morphogenesis of new particles occurs (V). During morphogenesis of mollivirus and pandoravirus, nuclear morphological changes are observed. The cycle ends with the release of progeny, which can occur by exocytosis or after cell lysis (VI). In this illustration, the pandoravirus was chosen in a representative way among the other groups of giant viruses. *Mimivirus viral factories organization is different from that illustrated in the figure, being more electron-dense than the cytoplasm and divided into at least two distinct areas.

Figure 4) Genome composition of giant viruses. Pie-chart representing the proportion of genes presenting no homology in databases (ORFans) and genes with the best hits with eukaryotes, prokaryotes and viruses for different groups of giant viruses. The data was considered at the time the first isolate for each group was described. Representatives include Cedratvirus A11, Pacmanvirus A23, Orpheovirus LCC2, Mollivirus sibericum, Marseillevirus T19, Pithovirus sibericum, Faustovirus E12, Kaumoebavirus isolate Sc and Pandoravirus salinus.

Figure 5) Phylogeny of NCLDV. Phylogenetic reconstruction based on amino acid sequences of DNA polymerase B family of different members of NCLDV group. Viral families containing viruses not associated with amoebae were

671 collapsed. Different colors represent distinct groups of amoebae giant viruses.
672 Blue: mimiviruses, green: pandoraviruses and mollivirus, red: marseilleviruses,
673 orange: pithoviruses, purple: asfarviridae-related viruses. Alignment was
674 performed using MUSCLE, and the tree was built using the Maximum
675 Likelihood method in FastTree software with 1,000 bootstrap replicates. Only
676 bootstrap values > 50 are shown. Scale bar represents rate of evolution.
677
678 **Figure 6) Virophage characteristics**. (A) Transmission electron microscopy
679 images of the Rio Negro virophage, a sputnik-related virus associated with the
680 Sambavirus (*Mimiviridae*), showing the virion shape within vesicles (left image),
681 and replication cycle inside viral factory (middle image), and effect on giant virus
682 particles (right image). White arrows indicate defective particles. VF: Viral
683 factory; (B) genome features of the Sputnik virus, the first virophage to be
684 discovered, indicating the genome size, ORF position, GC content and skew,
685 and the conserved genes found in members of the Lavidaviridae family; (C)
686 phylogenetic reconstruction based on MCP gene of virophages. Colors
687 represent the virophages identified by culture methods (red) and metagenomics
688 methods (blue). Alignment was performed using MUSCLE and the tree was
689 built using the Maximum Likelihood method in FastTree software with 1,000
690 bootstrap replicates. Only bootstrap values > 50 are shown. Scale bar
691 represents rate of evolution.

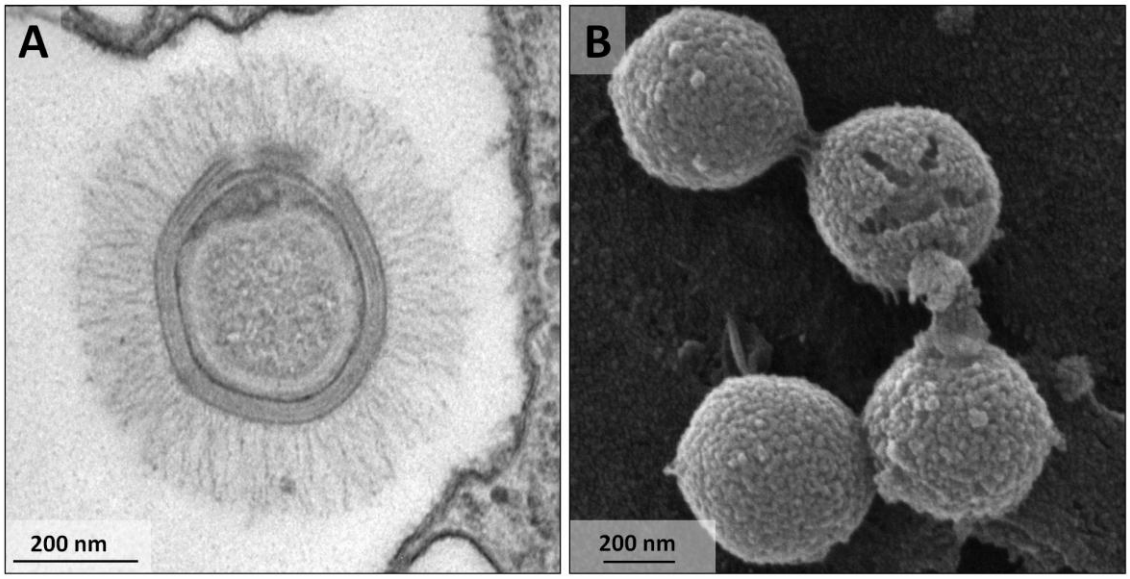


Figure 1) Electron microscopy images of mimivirus particles. (A) Transmission electron microscopy image demonstrating the capsid covered in fibrils and an internal membrane; (B) scanning electron microscopy image evidencing the stargate structure.

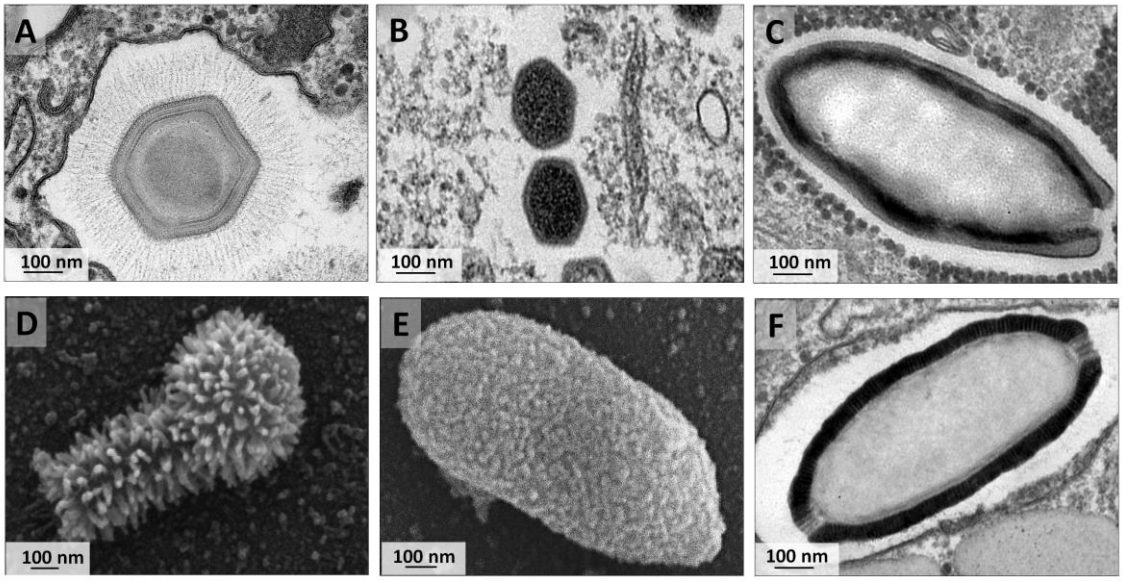


Figure 2) Diversity of giant viruses. Transmission and scanning electron microscopy images of different viral groups, representing the large variety of particle sizes and shapes among giant viruses. (A) mimivirus; (B) marseillevirus; (C) pandoravirus; (D) tupanvirus; (E) orpheovirus; (F) cedratvirus. Scale bars: 100 nm.

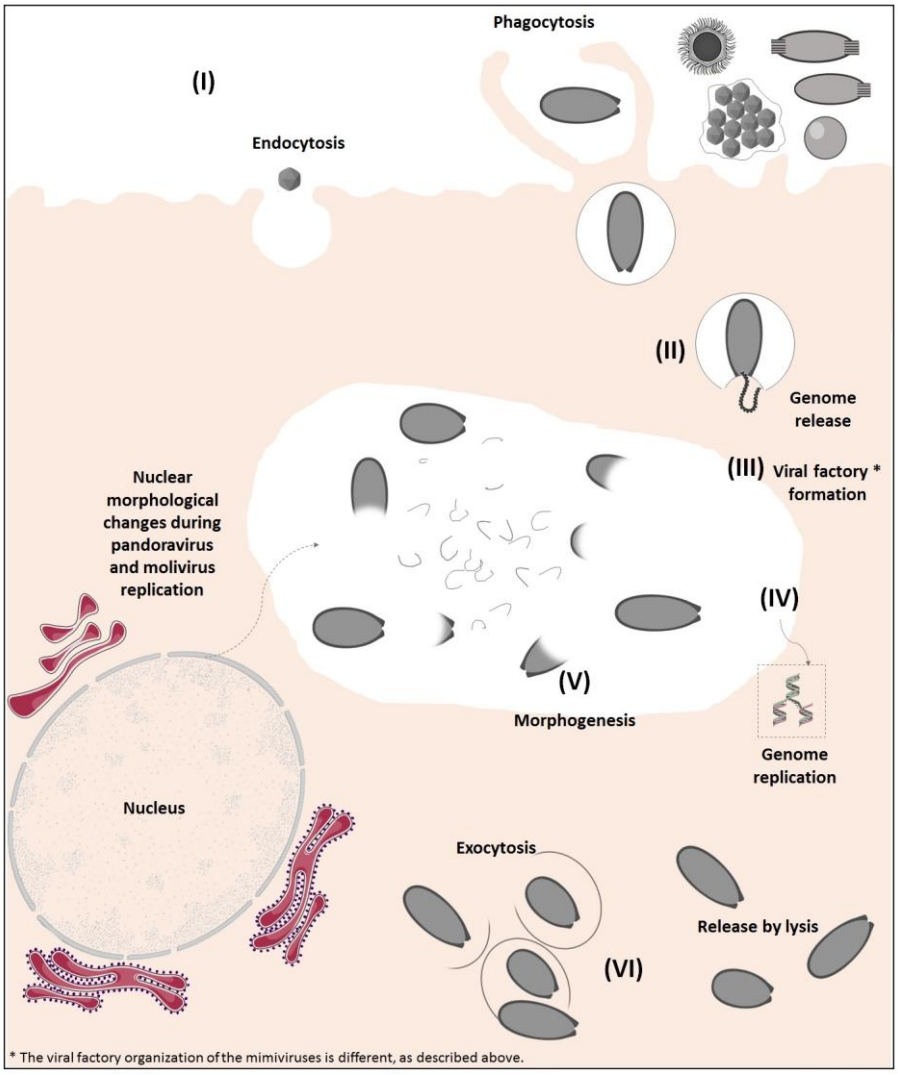


Figure 3: Representative scheme of the amoeba giant viruses replication cycle. The particles can enter host cell by phagocytosis or endocytosis for individual particles of marseillevirus (I). Subsequently, the genome is released into the host cytoplasm (II). A viral factory is established in the cytoplasm of the (III) wherein the genome replicates (IV) and the morphogenesis of new particles occurs (V). During morphogenesis of mollivirus and pandoravirus, nuclear morphological changes are observed. The cycle ends with the release of progeny, which can occur by exocytosis or after cell lysis (VI). In this illustration, the pandoravirus was chosen in a representative way among the other groups

of giant viruses. *Mimivirus viral factories organization is different from that illustrated in the figure, being more electron-dense than the cytoplasm and divided into at least two distinct areas.

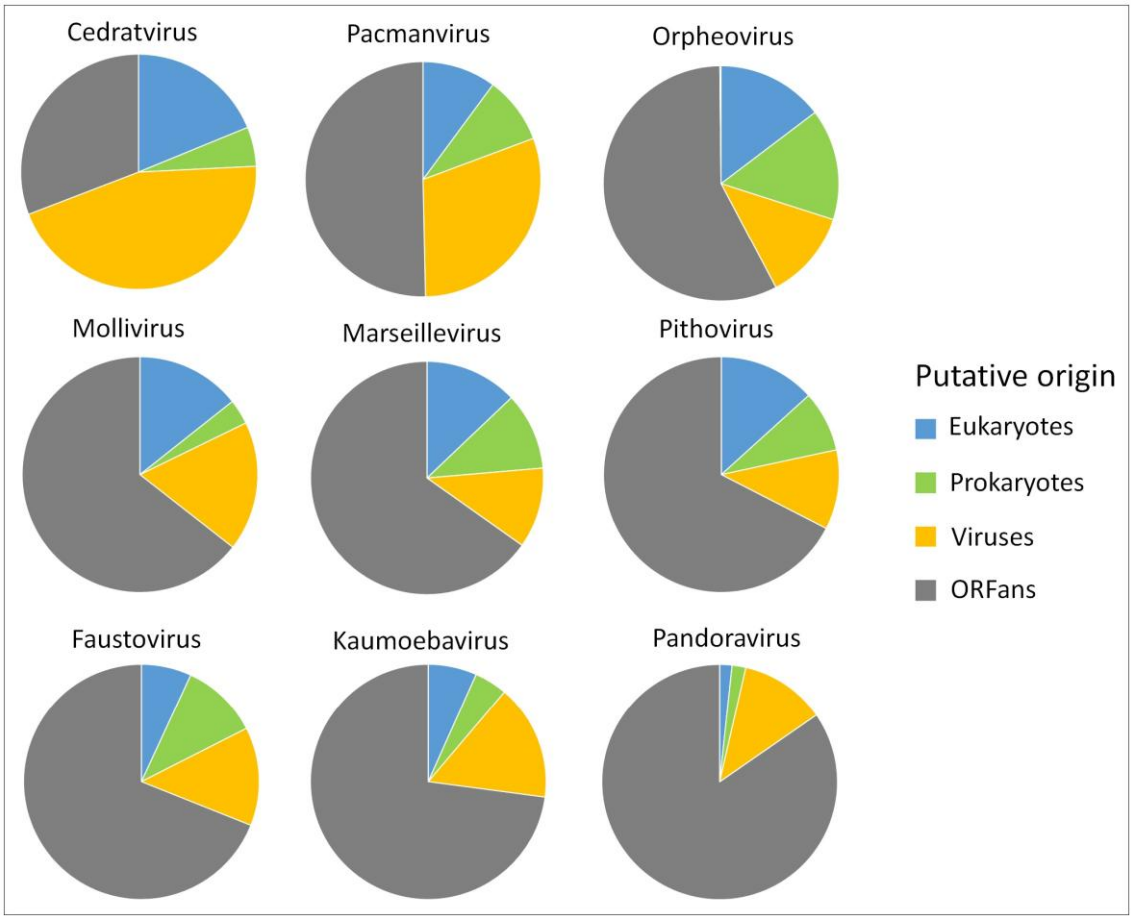


Figure 4) Genome composition of giant viruses. Pie-chart representing the proportion of genes presenting no homology in databases (ORFans) and genes with the best hits with eukaryotes, prokaryotes and viruses for different groups of giant viruses. The data was considered at the time the first isolate for each group was described. Representatives include Cedratvirus A11, Pacmanvirus A23, Orpheovirus LCC2, Mollivirus sibericum, Marseillevirus T19, Pithovirus sibericum, Faustovirus E12, Kaumoebavirus isolate Sc and Pandoravirus salinus.

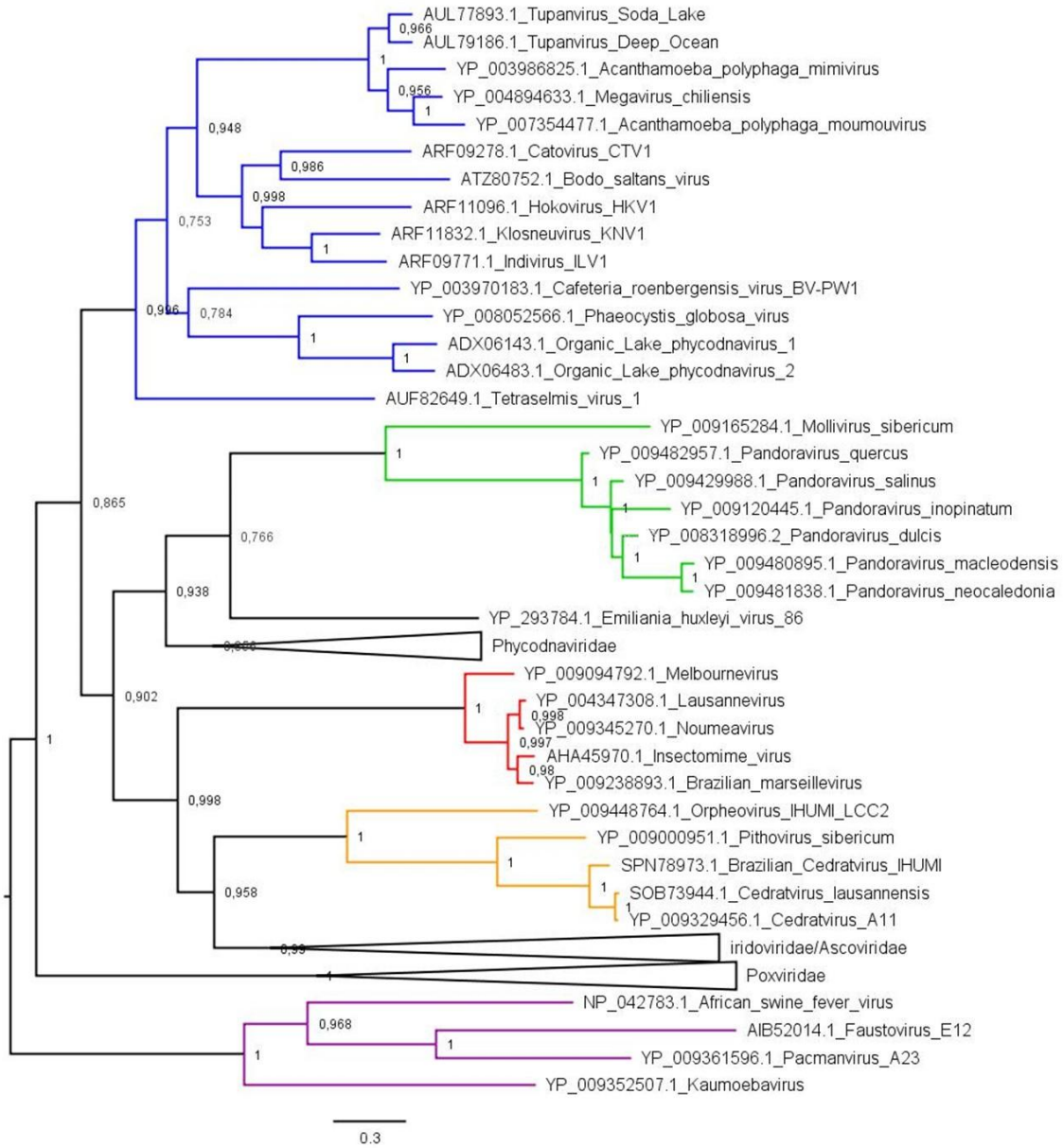


Figure 5) Phylogeny of NCLDV. Phylogenetic reconstruction based on amino acid sequences of DNA polymerase B family of different members of NCLDV group. Viral families containing viruses not associated with amoebae were collapsed. Different colors represent distinct groups of amoebae giant viruses. Blue: mimiviruses, green: pandoraviruses and mollivirus, red: marseilleviruses,

orange: pithoviruses, purple: asfarviridae-related viruses. Alignment was performed using MUSCLE, and the tree was built using the Maximum Likelihood method in FastTree software with 1,000 bootstrap replicates. Only bootstrap values > 50 are shown. Scale bar represents rate of evolution.

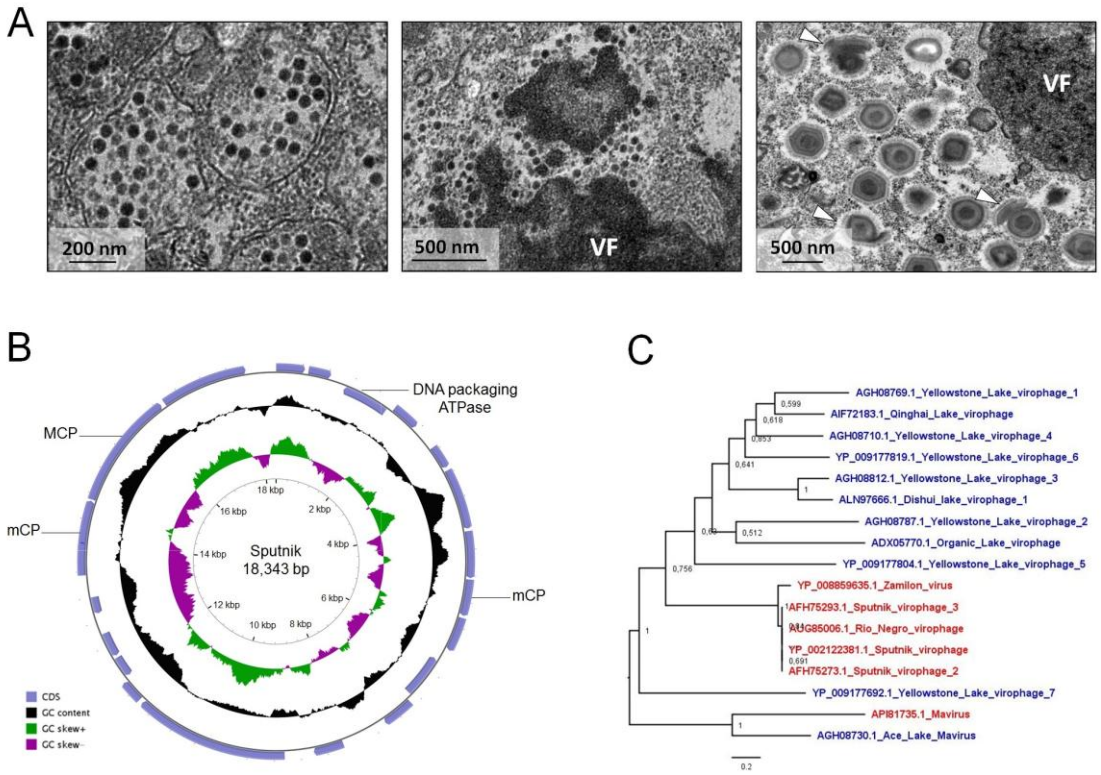


Figure 6) Virophage characteristics. (A) Transmission electron microscopy images of the Rio Negro virophage, a sputnik-related virus associated with the Sambavirus (*Mimiviridae*), showing the virion shape within vesicles (left image), and replication cycle inside viral factory (middle image), and effect on giant virus particles (right image). White arrows indicate defective particles. VF: Viral factory; (B) genome features of the Sputnik virus, the first virophage to be discovered, indicating the genome size, ORF position, GC content and skew, and the conserved genes found in members of the Lavidaviridae family; (C)

phylogenetic reconstruction based on MCP gene of virophages. Colors represent the virophages identified by culture methods (red) and metagenomics methods (blue). Alignment was performed using MUSCLE and the tree was built using the Maximum Likelihood method in FastTree software with 1,000 bootstrap replicates. Only bootstrap values > 50 are shown. Scale bar represents rate of evolution.