

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Microbiologia

Bruna Luiza de Azevedo

**DESCOBERTA E CARACTERIZAÇÃO DE ELEMENTOS GENÉTICOS MÓVEIS
ASSOCIADOS A UM VÍRUS GIGANTE**

Belo Horizonte

2021

Bruna Luiza de Azevedo

**DESCOBERTA E CARACTERIZAÇÃO DE ELEMENTOS GENÉTICOS MÓVEIS
ASSOCIADOS A UM VÍRUS GIGANTE**

Versão final

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Microbiologia.

Orientador: Dr. Jônatas Santos Abrahão

Coorientador: Dr. Rodrigo Araújo Lima Rodrigues

Belo Horizonte

2021

043

Azevedo, Bruna Luiza de.

Descoberta e caracterização de elementos genéticos móveis associados a um vírus gigante [manuscrito] / Bruna Luiza de Azevedo. – 2021.

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

Orientador: Dr. Jônatas Santos Abrahão. Coorientador: Dr. Rodrigo Araújo Lima Rodrigues.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. Mimiviridae. 3. Genoma. I. Abrahão, Jônatas Santos. II. Rodrigues, Rodrigo Araújo Lima. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 579



UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DA DEFESA DE DISSERTAÇÃO DE Bruna Luiza de Azevedo
Nº REGISTRO: 2019653618

Às 14 horas do dia 28 de abril de 2021, reuniu-se, na plataforma ZOOM - online, a Comissão Examinadora composta pelos Prof. Sávio Torres de Farias, Prof. Luiz Eduardo Del-Bem, o Prof. Jônatas Santos Abrahão- Orientador e o Dr. Rodrigo Araújo Lima Rodrigues - coorientador, para julgar o trabalho final "Descoberta e caracterização de elementos genéticos móveis associados a um vírus gigante", da aluna **Bruna Luiza de Azevedo**, requisito final para a obtenção do Grau de **MESTRE EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Jônatas Santos Abrahão, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. A candidata foi considerada **APROVADA**. O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. O candidato tem 60 (sessenta) dias, a partir desta data, para entregar a versão final da dissertação ao Programa de Pós-Graduação em Microbiologia da UFMG e requerer seu diploma.

Prof. Sávio Torres de Farias [titular externo à UFMG]

Prof. Luiz-Eduardo Del-Bem [titular interno à UFMG]

Prof. Jônatas Santos Abrahão [orientador]

Dr. Rodrigo Araújo Lima Rodrigues [coorientador]

Profª. Daniele da Glória de Souza
Coordenadora do PPGM/UFMG

*Ao meu amor, aos meus pais, ao Lucas e à
Nina. Por serem a base e o suporte de tudo*

AGRADECIMENTOS

Aos meus pais, por todo o seu esforço, amor, sabedoria, dedicação, paciência e cuidado. Sem vocês, sendo o meu suporte mais seguro, eu não chegaria nem perto de estar onde estou hoje.

Ao meu irmão, Lucas, meu companheiro, por ser uma pessoa tão incrível. Ao nascer você deixou a minha vida infinitamente mais feliz, trazendo leveza e amor.

Ao meu amor, Guilherme. Por todo o suporte, companheirismo, aprendizado e incentivo. Você me traz paz e amor nos momentos mais difíceis e me faz querer ser uma pessoa melhor a cada dia.

À minha filha pet, Nina. Mesmo que não seja alfabetizada na língua humana, merece toda a minha gratidão por representar um dos maiores pontos de equilíbrio e de paz na minha vida.

Ao meu orientador, Jônatas, por abrir as portas da virologia para mim, por confiar no meu trabalho, por nos presentear com projetos tão interessantes e desafiadores e por ser um professor tão dedicado e incrível. Você, com certeza, é uma inspiração para todos os seus alunos e alunas.

Ao meu coorientador, Rodrigo, pela infinita paciência e disponibilidade para me ajudar e me ensinar desde as coisas mais simples até as mais complexas. Esse trabalho é tão meu quanto seu.

Às minhas queridas amigas Isabella e Beatriz, minhas BBBs, que, desde o primeiro dia de graduação, estiveram ao meu lado para as mais diversas situações. Desde estudar, almoçar e jantar no bandeirão pagando um real e fofocar no banheiro do CAD, até se tornarem as minhas melhores amigas, que quero ao meu lado pelo resto da vida.

Às minhas migas, Isabella e Talita por tornarem minha passagem pelo Labvirus mais leve, divertida e segura.

A todos os colegas do GEPVIG, por formarem um grupo fantástico e inspirador e por estarem sempre disponíveis e abertos a me ensinar tantas coisas.

Aos colegas do Laboratório de Vírus, pela convivência e por transformarem o laboratório em um ambiente extremamente agradável, que representa a definição de parceria e de trabalho em grupo.

Aos professores do Laboratório de Vírus, Erna, Cláudio, Giliane e Betânia pela oportunidade de aprender com pesquisadores tão sábios, seja nas nossas reuniões ou em disciplinas.

Ao Centro de microscopia da UFMG, pelo trabalho extremamente bem qualificado, que enriquece as nossas pesquisas com imagens incríveis.

À banca avaliadora pela disponibilidade em colaborar para a melhoria deste trabalho.

Às agências de fomento, CNPq, CAPES e FAPEMIG pelo auxílio financeiro, que garantiu o desenvolvimento deste trabalho.

RESUMO

Os mimivírus são vírus gigantes de ameba, que podem ser infectados por vírus menores chamados de virófagos. Eles também possuem o seu próprio mobiloma, composto por diferentes elementos genéticos móveis, incluindo provirófagos e transpovirons. Os transpovirons são sequências móveis de DNA, flanqueadas por repetições terminais invertidas e separadas filogeneticamente em três clados correspondentes à linhagem do mimivírus ao qual estão associados (A, B ou C). Os virófagos são vírus que requerem tanto a ameba quanto o vírus gigante hospedeiro para se multiplicar e podem integrar seu genoma no DNA de seus hospedeiros, quando são chamados de provirófagos. Juntos, transpovirons e virófagos são importantes agentes de transferência gênica horizontal, impulsionando a evolução dos vírus gigantes. Entretanto, as pesquisas envolvendo esses elementos, ao redor do mundo, ainda são escassas. Neste trabalho descrevemos a descoberta e a caracterização genômica e filogenética de um novo isolado de mimivírus brasileiro, associado a um provirófago e a um transpoviron. Após sequenciamento genômico, montagem e predição de genes, identificamos o vírus gigante isolado como um mimivírus da linhagem A (1,2 Mb; 1.164 ORFs), denominado mimivírus *argentum*, bem como uma sequência de transpoviron (6.624 pb; 6 ORFs) e de um virófago (18.415 pb; 20 ORFs) associadas a ele. Análises posteriores indicaram a presença do genoma completo do virófago (um total de 20 ORFs) integrado ao genoma do mimivírus *argentum*. Ao nosso conhecimento, essa é a primeira evidência de um genoma completo de virófago integrado à sequência de um mimivírus. Os genes preditos para o virófago codificam proteínas conservadas da família *Lavidaviridae*, bem como proteínas possivelmente relacionadas com a integração em genomas hospedeiros. Já o genoma do transpoviron detectado apresenta três proteínas codificadas pelos *core genes* de transpovirons previamente descritos. No entanto, ele apresenta sequência e repetições terminais invertidas significativamente menores do que o que é descrito para os outros transpovirons conhecidos. O genoma do mimivírus *argentum* é similar ao dos outros mimivírus, codificando proteínas conservadas em *Nucleocyotoviricota* e, assim como o virófago e transpoviron, apresenta uma grande quantidade de proteínas hipotéticas, que são codificadas por genes que não apresentaram homólogos em bancos de dados no momento de sua descrição (ORFans). Por causa disso, comparamos a anotação realizada por meio do *BLASTp* e do *HHpred* e observamos que o uso do *HHpred* permitiu diminuir a quantidade de proteínas hipotéticas/ORFans que seriam anotadas para os três genomas estudados. Com o *HHpred*, também conseguimos prever funções e/ou estruturas para proteínas de outros transpovirons descritos que estavam anotadas como hipotéticas. Os transpovirons conhecidos codificam um conjunto de 10 proteínas diferentes, sendo que o clado B apresenta maior diversidade de proteínas. As diferentes análises comparativas entre genomas dos transpovirons conhecidos, bem como as análises filogenéticas sugerem que o clado A de transpovirons seja o mais conservado dentre os outros três clados propostos e que a relação entre os transpovirons e mimivírus pode ter se estabelecido antes da divergência das linhagens de mimivírus. Em conjunto, os resultados obtidos expandem o conhecimento básico sobre elementos ainda pouco estudados e estudos futuros, sobretudo de caráter biológico, podem contribuir para o melhor entendimento da relação entre os vírus gigantes, seu mobiloma e a evolução.

Palavras-chave: vírus gigantes, mimivírus, virófago, transpoviron, mobiloma, provirófago.

ABSTRACT

Mimiviruses are giant viruses of amoeba that can be infected by smaller viruses denominated virophages. They also possess their own mobilome, composed by different mobile genetic elements which include provirophages and transpovirons. Transpovirons are mobile DNA sequences flanked by terminal inverted repeats and phylogenetically separated in three clades corresponding to the lineage of mimivirus to which they are associated (A, B, or C). Virophages are viruses that require an amoeba and a giant virus to replicate. They can integrate their sequence into their hosts genomes forming provirophages. Together, transpovirons and virophages are important agents of horizontal gene transfer and promote giant viruses' evolution. However, studies that involve these elements are still scarce around the world. In this work, we performed a genomic and phylogenetic characterization of a new Brazilian mimivirus isolate, associated to a provirophage and a transpoviron. After genomic sequencing, assembly, and genes prediction we identified the new isolate as a lineage A mimivirus (1.2 Mb; 1,164 ORFs), called mimivirus argentum, as well as two associated sequences: a transpoviron (6,624 bp; 6 ORFs) and a virophage (18,415 bp; 20 ORFs). Further analysis indicated the presence of the full virophage genome (a total of 20 ORFs) integrated into the mimivirus argentum genome. To our knowledge this is the first evidence of a virophage full genome integrated into a mimivirus sequence. The predicted genes for the virophage code the *Lavidaviridae* family conserved proteins, as well as proteins possibly related to integration into host genomes. The genome of detected transpoviron presents three proteins encoded by core genes from previously described transpovirons. However, its sequence and terminal inverted repeats are significantly smaller than those described for other known transpovirons. Regarding mimivirus argentum genome, it is like those from other mimiviruses and code *Nucleocyotoviricota* conserved proteins. As well as the transpoviron and the virophage, mimivirus argentum presents a great proportion of hypothetical proteins, which are encoded by genes that did not present homologs in databases at the time of its description (ORFans). Based on this, we compared the annotation made using BLASTp and HHpred and we observed that the use of HHpred were able to decrease the number of hypothetical proteins/ORFans that would be annotated to the three studied genomes. By using HHpred, we also predicted functions and/or structures to proteins annotated as hypothetical from other transpovirons already described. The known transpovirons encode a set of 10 different proteins, with clade B presenting a greater diversity of proteins. The different comparative analysis between known transpovirons genomes, as well as phylogeny, suggest that clade A transpovirons are most conserved than the three other proposed clades and that the relationship between transpovirons and mimivirus may have been established before the divergence of mimiviruses lineages. Altogether our results expand the basic knowledge about elements poorly studied. Future studies, especially those of biological nature, may contribute to a better understanding of the relationship between giant viruses, their mobilome and evolution.

Keywords: giant virus, mimivirus, virophage, transpoviron, mobilome, provirophage

LISTA DE FIGURAS

Figura 1 - Imagem de microscopia óptica de ameba após coloração de Gram.....	15
Figura 2 – Diversidade morfológica de vírus gigantes de ameba, membros do grupo dos NCLDV.....	19
Figura 3 - Os dois vírus de ameba brasileiros que chamaram a atenção da comunidade científica.....	21
Figura 4 - Esquema representando a hierarquia taxonômica atual dos NCLDV.....	22
Figura 5 – Os diferentes clados da família <i>Mimiviridae</i>	24
Figura 6 – Aspectos estruturais da partícula de um mimivírus.....	25
Figura 7 - Esquema indicando a taxonomia atual dos virófagos.....	34
Figura 8 - Imagem de microscopia eletrônica de transmissão de partículas do virófago sputnik.....	40
Figura 9 - Esquema que resume a estrutura geral de partícula do virófago sputnik.....	41
Figura 10 - Esquema representativo dos diferentes passos da coinfeção de uma célula de ameba pelo virófago sputnik e por um mimivírus.....	43
Figura 11 - Esquema ilustrando os níveis mais altos da classificação hierárquica proposta por Wicker e colaboradores (2007) para os transposons.....	50
Figura 12 - Árvore filogenética de sequências concatenadas das proteínas codificadas pelos <i>core genes</i> dos transpovirons.....	53
Figura 13 - Imagens de microscopia eletrônica de transmissão mostrando uma visão geral dos novos isolados de mimivírus e de virófago dentro da ameba.....	67
Figura 14 - Caracterização dos <i>best hits</i> obtidos no <i>BLASTp</i> , durante análise das proteínas do mimivírus <i>argentum</i> , em busca por similaridade em bancos de dados.....	69
Figura 15 - Resultados das análises no <i>HHpred</i> para as proteínas do mimivírus <i>argentum</i> , cujos <i>best hits</i> do <i>BLASTp</i> eram proteínas hipotéticas (ORFans).....	71
Figura 16 - Comparação entre os dados obtidos após a análise do conjunto de proteínas do mimivírus <i>argentum</i> no <i>BLASTp</i> e no <i>HHpred</i>	73
Figura 17 - Mapa do genoma do mimivírus <i>argentum</i> , indicando características como a distribuição das ORFs, conteúdo de G-C e RNAs transportadores na sequência.....	75
Figura 18 - Filogenia de <i>Nucleocyotoviricota</i> , com base na sequência de aminoácidos da enzima DNA polimerase B, presente no mimivírus <i>argentum</i> (M731)	76
Figura 19 - Caracterização do genoma do virófago sputnik <i>argentum</i>	81

Figura 20 -Filogenia das proteínas S7, S6, S11 e S17 do virófago sputnik argentum.....	83
Figura 21 - Gráfico de pontos (<i>Dot plot</i>) gerado após alinhamento das sequências do virófago sputnik argentum e do mimivírus argentum, utilizando o <i>BLASTn</i>	85
Figura 22 – Mapa do genoma do mimivírus argentum, destacando a possível região de integração da sequência do virófago sputnik argentum.....	87
Figura 23 - Número de proteínas contendo repetição de tripla hélice de colágeno em alguns diferentes mimivírus.....	88
Figura 24 – Análise de sintenia dos genes para as proteínas contendo repetição de tripla hélice de colágeno presentes em diferentes virófagos do gênero <i>Sputnikvirus</i>	89
Figura 25 - Filogenia de uma das proteínas contendo repetição de tripla hélice de colágeno (S14) do virófago sputnik argentum.....	91
Figura 26 – Caracterização genômica do transpoviron do mimivírus argentum (Tvmarg).....	94
Figura 27 – Mapas individuais dos genomas dos transpovirons descritos, indicando a sintenia dos genes e as respectivas proteínas codificadas por cada um deles.....	99
Figura 28 - Diagrama de Venn que resume a ocorrência das proteínas de transpovirons em cada um dos seus três clados (A, B e C)	100
Figura 29 - Filogenia das proteínas T1, T5 e T6 do transpoviron do mimivírus argentum.....	101
Figura 30 - Filogenia das proteínas T2 e T4 do transpoviron do mimivírus argentum.....	102

LISTA DE TABELAS

Tabela 1 – Características gerais do genoma de alguns mimivírus das linhagens A, B e C.....	28
Tabela 2 – Lista de virófagos descritos e algumas de suas principais características.....	37
Tabela 3 – Resumo dos seis transpovirons já descritos e os seus mimivírus associados.....	52
Tabela 4 - Caracterização no <i>BLASTn</i> das três diferentes sequências estudadas neste trabalho, obtidas após o sequenciamento e montagem.....	68
Tabela 5 – Diferenças entre <i>best hits</i> obtidos no <i>BLASTp</i> e no <i>HHpred</i> para algumas proteínas do mimivírus <i>argentum</i>	74
Tabela 6 – Resultados da análise das ORFs preditas para o virófago <i>sputnik argentum</i> no <i>BLASTp</i>	77
Tabela 7 – Análise das proteínas do <i>sputnik argentum</i> no <i>HHpred</i> , em comparação com os domínios e funções preditos anteriormente para as suas sequências homólogas.....	79
Tabela 8 – Resultados do alinhamento (<i>BLASTn</i>) da sequência de nucleotídeos do <i>sputnik argentum</i> contra outros virófagos isolados e comparação de suas características genômicas....	82
Tabela 9 - Resultados do alinhamento (<i>BLASTn</i>) entre a sequência do virófago <i>sputnik 2</i> contra o genoma do <i>lentillevírus</i> , fragmentado em 10 diferentes <i>contigs</i>	86
Tabela 10 – Resultados da comparação, realizada no <i>BLASTn</i> , entre as proteínas contendo repetição de tripla hélice de colágeno de virófagos do gênero <i>Sputnikvirus</i>	90
Tabela 11 – Análise das proteínas do transpoviron do mimivírus <i>argentum</i> utilizando o <i>BLASTp</i>	92
Tabela 12 – Dados da busca por domínios e sequências homólogas para as proteínas do transpoviron do mimivírus <i>argentum</i> realizadas no <i>InterProScan</i> e no <i>HHpred</i>	93
Tabela 13 – Resumo destacando as proteínas hipotéticas de transpovirons que apresentaram resultados diferentes da sua anotação, após busca similaridade no <i>HHpred</i>	96
Tabela 14 – Resultados do alinhamento (<i>BLASTn</i>) da sequência de nucleotídeos do transpoviron do mimivírus <i>argentum</i> contra os outros transpovirons descritos e comparação de suas características genômicas.....	97

LISTA DE SIGLAS E ABREVIATURAS

μL – Microlitro

μm – Micrômetros

aaRS - aminoacil-tRNA sintetase

ACMV – Acanthamoeba castellanii mamavírus

ALM – Ace Lake Mavírus

APMV – Acanthamoeba polyphaga mimivírus

ArgRS - Arginil-tRNA sintetase

A-T – Adenina-Timina

CEOs - *Capsid-encoding organisms* (Organismos codificadores de capsídeo)

cm² - Centímetros quadrados

CM-UFMG – Centro de Microscopia da Universidade Federal de Minas Gerais

CroV – Cafeteria roenbergensis virus

CysRS - Cisteinil-tRNA sintetase

DJR – *Double-jelly-roll*

DLV -Dishui Lake Virophage

DNAfd – DNA fita-dupla

EGMs – Elementos Genéticos Móveis

ETs – Elementos Transponíveis

FV - Fábrica viral

G-C – Guanina-Citosina

GEPVIG – Grupo de Estudo e Prospecção de Vírus Gigantes

h.p.i – Horas pós-infecção

ICTV - *International Committee on Taxonomy of Viruses* (Comitê Internacional em Taxonomia de Vírus)

kpb – Quilopares de base

LRR – *Leucine-rich repeat* (Repetição rica em leucina)

LTR – *Long Terminal Repeats* (Repetições Terminais Longas)

M.O.I – *Multiplicity of infection* (Multiplicidade de infecção)

Mb – Megabases

MCP – *Major Capsid Protein* (Proteína principal do capsídeo)

mCP – *Minor Capsid Protein* (Proteína secundária do capsídeo)

MET – Microscopia eletrônica de transmissão

MetRS - Metionil-tRNA sintetase

mg – Miligrama

mL – Mililitro

NCLDV – *Nucleo-Cytoplasmic Large DNA Viruses* (Vírus Grandes Núcleo- Citoplasmáticos de DNA)

nm - Nanômetros

nt – Nucleotídeos

nt – Nucleotídeos

NYMV - Niemeyer vírus

OLV – Organic Lake Virophage

ORFs – *Open Reading Frames* (Fases abertas de leitura)

p/v – Peso/volume

pb – Pares de bases

PBS – Tampão fosfato salino

PgV-16T: Phaeocystis globosa virus -16T

PgVV - Phaeocystis globosa virus virophage

PYG – *Peptone-Yeast extract-Glucose* (Peptona-Extrato de levedura-Glicose)

QLV -Qinghai Lake Virophage

REOs - *Ribosome-encoding organisms* (Organismos codificadores de ribossomo)

RNA_m – RNA mensageiro

RNA_r – RNA ribossomal

RNA_t – RNA transportador

rpm – Rotações por minuto

TGH – Transferência Gênica Horizontal

TIR – *Terminal Inverted Repeats* (Repetições Terminais Invertidas)

TPR - *Tetratricopeptide repeat* (Repetição tetratricopeptídeo)

Tvcourdo - Transpoviron do megavírus courdo7

Tvlent - Transpoviron do lentillevírus

Tvmama - Transpoviron do mamavírus

Tvmarg – Transpoviron do mimivírus argenteum

Tvmaust - Transpoviron do moutovírus australiensis

Tvmonve - Transpoviron do moutovírus monve

Tvmvitis - Transpoviron do megavírus vitis

TyrRS - Tirozil-rRNA sintetase

VLEs – *Virophage-like elements* (Elementos similares a virófagos)

YSLV – YellowStone Lake Virophage

SUMÁRIO

1	INTRODUÇÃO	17
1.1	A descoberta dos vírus gigantes e seu impacto para a virologia.....	17
1.2	Vírus Grandes Núcleo-Citoplasmáticos de DNA (NCLDV): um grupo em expansão	20
1.3	Os mimivírus	26
1.3.1	A família <i>Mimiviridae</i>	26
1.3.2	Estrutura da partícula do mimivírus	27
1.3.3	Ciclo de multiplicação do mimivírus	29
1.3.4	Genoma dos mimivírus.....	30
1.4	Os virófagos.....	35
1.4.1	A descoberta dos virófagos	35
1.4.2	A família <i>Lavidaviridae</i>	37
1.5	Os virófagos sputnik.....	42
1.5.1	Estrutura da partícula.....	42
1.5.2	Ciclo de multiplicação dos virófagos sputnik	44
1.5.3	Genoma dos virófagos sputnik	46
1.6	Os elementos genéticos móveis.....	49
1.6.1	Transposons	50
1.7	Vírus gigantes e seu mobiloma.....	53
1.7.1	Provirófagos	54
1.7.2	Transpovirons	55
1.8	O papel do mobiloma na evolução dos vírus gigantes e virófagos	57
2	JUSTIFICATIVA	59
3	OBJETIVOS	60
4	FLUXOGRAMA	61
5	MATERIAIS E MÉTODOS	62
5.1	Isolamento viral e obtenção de imagens de microscopia dos isolados.....	62
5.1.1	Cultivo celular	62
5.1.2	Isolamento viral	62
5.1.3	Multiplicação e purificação viral.....	63
5.1.4	Titulação viral.....	64
5.1.5	Microscopia eletrônica de transmissão.....	64

5.2 Caracterização genômica dos isolados e do transpoviron	65
5.2.1 Sequenciamento, montagem e predição de genes	65
5.2.2 Anotação de genes	66
5.2.3 Análises genômicas comparativas	67
5.3 Análises filogenéticas	67
6 RESULTADOS	69
6.1 Isolamento e imagens de microscopia eletrônica dos isolados.....	69
6.2 Sequenciamento e montagem dos genomas	70
6.3 Caracterização genômica do mimivirus argenteum	71
6.3.1 Anotação dos genes	71
6.3.2 Filogenia do mimivirus argenteum	78
6.4 Caracterização genômica do sputnik argenteum	80
6.4.1 Anotação dos genes	80
6.4.2 Filogenia do sputnik argenteum	85
6.5 Mimivirus argenteum e seu mobiloma.....	87
6.5.1 Detecção do provirófago do sputnik argenteum no genoma do mimivirus argenteum	87
6.5.2 Análise comparativa das proteínas contendo repetição de tripla-hélice de colágeno presentes em viróforos e vírus gigantes.	90
6.5.3 Caracterização do transpoviron do mimivirus argenteum e comparação com os outros transpovirons descritos.	95
6.5.3.1 Genoma do transpoviron do mimivirus argenteum.....	95
6.5.3.2 Análise das proteínas de todos os transpovirons descritos por meio do <i>HHpred</i>	98
6.5.3.3 Comparação entre o transpoviron do mimivirus argenteum e os outros transpovirons descritos.....	99
6.5.3.4 Análises filogenéticas das proteínas do transpoviron do mimivirus argenteum	103
7 DISCUSSÃO	107
8 CONCLUSÕES	114
REFERÊNCIAS BIBLIOGRÁFICAS	116
ANEXOS	128
Anexo 1 - Protocolo de preparação do meio de cultura peptona/extrato de levedura/glicose (PYG).	128
Anexo 2 – <i>Best hits</i> obtidos para cada proteína do mimivirus argenteum no <i>BLASTp</i> e no <i>HHpred</i>	129
Anexo 3 – <i>Best hits</i> para proteínas dos transpovirons já conhecidos anteriormente, obtidos no <i>HHpred</i>	149

Anexo 4 - Alinhamento (<i>BLASTp</i>) das proteínas do transpoviron do mimivírus argenteus contra o conjunto de proteínas dos outros transpovirons já descritos.	150
---	-----

1 INTRODUÇÃO

1.1 A descoberta dos vírus gigantes e seu impacto para a virologia

Os vírus são entidades biológicas muito diversas e abundantes no planeta Terra, que são capazes de utilizar diferentes formas de vida como hospedeiros (KOONIN *et al.*, 2020). Eles são definidos como parasitas intracelulares obrigatórios, dependentes da maquinaria metabólica das suas células hospedeiras. Eles são, ainda, frequentemente definidos por conter apenas um tipo de ácido nucleico em suas partículas e costumam ser diferenciados de outros microrganismos, como as bactérias, com base no tamanho reduzido de seus genomas e de suas partículas (LWOFF, 1957; RAOULT E FORTERRE, 2008). Por serem tão pequenos, acreditava-se que os vírus não podiam ser visualizados por meio de microscópios ópticos e que eram sempre agentes filtráveis, ou seja, capazes de passar por filtros com poros de 0,2 μm (LWOFF, 1957; RAOULT E FORTERRE, 2008). Dessa forma, os métodos utilizados para identificação e isolamento de novos vírus acabavam impedindo a descoberta de vírus com fenótipos diferentes do proposto. (COLSON *et al.*, 2018).

Grande parte dessas definições começaram a ser repensadas quando, em 2003, foi apresentado à comunidade científica o primeiro vírus gigante (LA SCOLA *et al.*, 2003). A história da descoberta dos vírus gigantes começou em 1992, quando na tentativa de descobrir o agente causador de um surto de pneumonia na cidade inglesa de Bradford, pesquisadores coletaram amostras de água da torre de refrigeração de um hospital e isolaram um organismo intracelular, utilizando amebas de vida livre da espécie *Acanthamoeba polyphaga* como plataforma. As análises iniciais dessas amostras, utilizando microscopia óptica e coloração de Gram, levaram os pesquisadores a concluir que o novo isolado era uma pequena bactéria Gram-positiva com morfologia de cocos, que foi nomeada provisoriamente de cocos de Bradford (Figura 1) (LA SCOLA *et al.*, 2003; ROLLAND *et al.*, 2019).

Alguns anos mais tarde essas amostras foram reanalisadas, em parceria com cientistas franceses, com o objetivo de identificar os cocos de Bradford. Durante esses estudos, os pesquisadores resolveram analisar as amostras utilizando a microscopia eletrônica, uma vez que as várias tentativas de extração do material genético e de detecção do gene para o RNA ribossomal (RNAr) 16S bacteriano falharam. Surpreendentemente, as imagens de microscopia mostraram, ao invés de bactérias, estruturas icosaédricas típicas de vírus de, aproximadamente,

600 nanômetros (nm) (LA SCOLA *et al.*, 2003; ROLLAND *et al.*, 2019) Além disso, eles também observaram que o ciclo de multiplicação desse organismo apresentava uma fase de eclipse, que é uma característica típica de vírus (LA SCOLA *et al.*, 2003). A natureza viral desse microrganismo foi confirmada com o sequenciamento do seu genoma de quase 1,2 milhão de pares bases (pb), número muito elevado se comparado aos genomas dos vírus conhecidos até então (RAOULT *et al.*, 2004). O novo vírus foi chamado de *Acanthamoeba polyphaga* mimivirus (APMV), ou simplesmente mimivírus, já que é capaz de infectar a ameba *Acanthamoeba polyphaga* e de mimetizar uma bactéria, devido ao seu tamanho e à coloração de Gram (LA SCOLA *et al.*, 2003). A descoberta de um vírus com características tão peculiares trouxe de volta o debate acerca do conceito de vírus aceito até então.

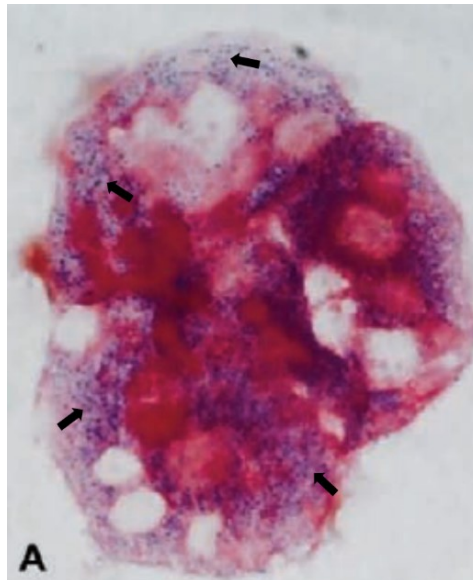


Figura 1 - Imagem de microscopia óptica de ameba após coloração de Gram. É possível observar, dentro da célula, os chamados cocos de Bradford (mais tarde identificados como mimivírus), corados em roxo (setas pretas). **Fonte:** Modificado de RAOULT; LA SCOLA; BIRTLES, 2007, p.97

Graças à complexidade e ao tamanho de sua partícula, os mimivírus contradizem a ideia de que todos os vírus são menores do que organismos celulares, já que seu tamanho é comparável a bactérias do grupo dos micoplasmas (RAOULT *et al.*, 2004). Essa característica permite a visualização desses vírus por microscopia óptica, além de permitir que eles fiquem retidos nos filtros com poros de 0,2 μm (LA SCOLA *et al.*, 2003). Além de uma grande partícula, os mimivírus também apresentam um grande e complexo genoma de DNA fita-dupla (DNAfd), carregando, por consequência, uma expressiva quantidade de genes que nunca havia sido descrita em outros vírus. Dentre eles, pode-se destacar um importante conjunto de genes necessários em várias etapas da tradução de proteínas (RAOULT *et al.*, 2004). Dessa forma, a

ideia de que os vírus não possuem elementos relacionados ao processo traducional e de que seus genomas são pequenos e pouco complexos também começou a ser desconstruída. Outro ponto interessante descoberto nas análises iniciais do genoma e da partícula dos mimivírus foi a detecção de transcritos virais associados às partículas do APMV, ou seja, moléculas de RNA mensageiro (RNAm) sendo carregados pelas partículas infecciosas de mimivírus (RAOULT *et al.*, 2004). Apesar de a presença de DNA e RNA empacotados em uma mesma partícula viral já ter sido descrita em outros vírus grandes, como os da família *Herpesviridae*, a descoberta dessa característica nos mimivírus contribuiu para contrapor a ideia de que os vírus carregam apenas um tipo de ácido nucleico em suas partículas (RAOULT *et al.*, 2004).

Além da definição do conceito de vírus, a descoberta dos vírus gigantes também reacendeu o debate que envolve a separação clássica feita entre os vírus e os organismos celulares. Na década de 1990, Woese e colaboradores propuseram a divisão de todos os organismos vivos em três domínios: Bacteria, Archaea e Eukarya. Essa divisão foi baseada, principalmente, em sequências de RNAr (RNA ribossomal), que estão presentes nos três grupos propostos (WOESE; KANDLER; WHEELIS, 1990). Alguns trabalhos posteriores à descoberta dos mimivírus começaram a questionar essa divisão, alegando que o agrupamento de todos os organismos em uma mesma árvore filogenética, com base em sequências como o RNAr, naturalmente excluiria os vírus, já que não há evidências de que eles possuam ribossomos. Em contrapartida, um outro conjunto de genes, compartilhado tanto pelos organismos celulares quanto pelos vírus gigantes poderia ser utilizado para tal classificação (RAOULT; FORTERRE, 2008; BOYER *et al.*, 2010).

Em consonância com essa discussão, Raoult e Forterre (2008) propuseram uma classificação dicotômica, não natural, que incluiria tanto os vírus, quanto os organismos celulares. Dessa forma, os componentes dos domínios Bacteria, Archaea e Eukarya formariam o grupo dos organismos codificadores de ribossomos (REOs, *ribosome-encoding organisms*), enquanto os vírus fariam parte do grupo dos organismos codificadores de capsídeo (CEOs, *capsid-encoding organisms*) (RAOULT; FORTERRE, 2008). Com base nessa nova proposta de classificação dos organismos e na redefinição dos conceitos atribuídos aos vírus, foi proposto que eles poderiam ser definidos como: “um CEO, que é composto por proteínas e ácidos nucleicos, que se monta em um nucleocapsídeo e que usa um REO para concluir o seu ciclo de vida” (RAOULT; FORTERRE, 2008, tradução nossa). No entanto, com a expansão do grupo de vírus gigantes ocorrida nos últimos anos, novas informações vêm sendo obtidas, fazendo com que esse tema ainda permaneça em debate e em constante atualização.

1.2 Vírus Grandes Núcleo-Citoplasmáticos de DNA (NCLDV): um grupo em expansão

Ao analisar o conjunto de todos os vírus conhecidos, não é possível identificar genes universais ou conservados entre todos eles, como acontece com os organismos celulares. Esse fato contribui com a ideia de que os vírus formem um grupo polifilético, tendo múltiplas origens evolutivas (KOONIN *et al.*, 2020). Em contrapartida, a análise de alguns grupos isolados de vírus sugere que eles possam compartilhar seu próprio ancestral comum. À exemplo disso, pode-se citar o grupo dos Vírus Grandes Núcleo-Citoplasmáticos de DNA (NCLDV – *Nucleo-Cytoplasmic Large DNA Viruses*), que é um grupo sugerido como monofilético e que compartilha um conjunto de genes codificadores de proteínas homólogas (IYER; ARAVIND; KOONIN, 2001).

De maneira geral, os NCLDV são caracterizados por serem vírus com grandes partículas, geralmente maiores do que 150 nm, e grandes genomas de DNAfd, maiores do que 100 quilo pares de bases (kpb). Eles infectam uma ampla gama de hospedeiros eucariotos, que engloba desde protistas, como algas e amebas, até animais, como insetos e mamíferos. (FILÉE; CHANDLER, 2010; COLSON *et al.*, 2012). Eles também compartilham a capacidade de completar o seu ciclo de multiplicação inteiramente no citoplasma hospedeiro ou de iniciarem o seu ciclo no núcleo celular, terminando no citoplasma. Isso sugere que eles possam apresentar uma menor dependência da maquinaria de transcrição e replicação da célula hospedeira (IYER *et al.*, 2006). Em contrapartida, vírus que não apresentam algumas dessas características generalistas e mesmo assim se associam filogeneticamente ao grupo vêm sendo descritos recentemente (SUBRAMANIAM *et al.*, 2020)

Embora já descrito e estudado anteriormente, o grupo dos NCLDV ganhou maior destaque quando os recém-descobertos mimivírus foram filogeneticamente agrupados em uma linhagem divergente dentro desse grupo, formando um novo táxon, a família *Mimiviridae* (LA SCOLA *et al.*, 2003; IYER *et al.*, 2006). Dessa forma, a adição de um tipo tão diferente de vírus ao grupo aumentou substancialmente o interesse pela busca por novos vírus e por pesquisas aprofundadas sobre vários aspectos dos NCLDV, como a sua origem, evolução e taxonomia. Antes da descrição da família *Mimiviridae*, os NCLDV já abrigavam as famílias *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae* e *Phycodnaviridae*, com base em um conjunto de genes compartilhados por elas (IYER; ARAVIND; KOONIN, 2001).

Ao longo dos últimos anos, novos vírus gigantes de ameba foram sendo descobertos, graças ao desenvolvimento de novas técnicas específicas para o isolamento desses organismos em

Acanthamoeba spp. (PAGNIER *et al.*, 2013; KHALIL; ANDREANI; LA SCOLA, 2016). Essas inúmeras descobertas têm levado, por consequência, a uma expansão do grupo dos NCLDV. Além de diferentes linhagens da família *Mimiviridae*, outros vírus gigantes de ameba muito diferentes e peculiares foram descobertos, elucidando a diversidade existente no grupo (Figura 2). O primeiro *Marseillevirus*, por exemplo, isolado na França em 2009, é um vírus menor do que o APMV. Suas partículas apresentam um tamanho aproximado de 250 nm de diâmetro (Figura 2 B) e o seu genoma é composto por 368.454 pb, codificando cerca de 457 proteínas. A análise do genoma desse vírus mostrou que ele possui um conjunto de 28 genes que estão presentes também nos NCLDV. Somadas a isso, as análises filogenéticas sugeriram que os *marseillevirus* fizessem parte de uma nova família viral dentro do grupo, a família *Marseilleviridae* (BOYER *et al.*, 2009). Um outro grupo de vírus de *Acanthamoeba* spp. com partículas icosaédricas menores do que o APMV (~260 nm) são os medusavírus. Eles foram isolados de amostras de água termal do Japão e suas partículas apresentam pequenas estruturas “spike” na superfície do capsídeo. Além disso, esses vírus parecem utilizar o núcleo da ameba hospedeira para a replicação do seu DNA (YOSHIKAWA *et al.*, 2019).

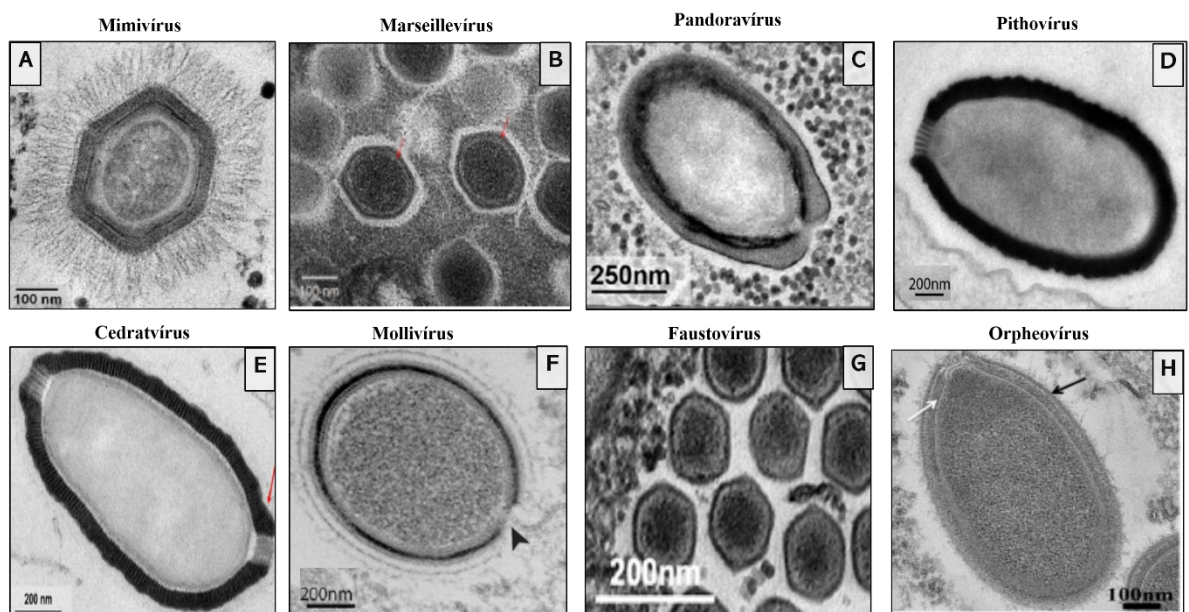


Figura 2 – Diversidade morfológica de vírus gigantes de ameba, membros do grupo dos NCLDV. (A-H) Imagens de Microscopia Eletrônica de Transmissão (MET) mostrando partículas virais de mimivírus, *Marseillevirus*, pandoravírus, pithovírus, cedratvírus, mollivírus, Faustovírus e orpheovírus, respectivamente. **Fonte:** Modificado de RODRIGUES; SILVA; ABRAHÃO, 2020; ARANTES *et al.*, 2016; ANDRADE *et al.*, 2019; ABERGEL; LEGENDRE; CLAVERIE, 2015; SILVA *et al.*, 2018.; LEGENDRE *et al.*, 2015; de BORGES *et al.*, 2019; ANDREANI *et al.*, 2018.

Por algum tempo depois de sua descoberta, os mimivírus foram donos do posto de vírus com a maior partícula e genoma da virosfera. No entanto, os pandoravírus, descritos em 2013, foram os primeiros a tomar esse posto, já que apresentam partículas ovoides que podem alcançar 1 µm de comprimento (Figura 2 C), além de genomas que podem conter 2,5 megabases (Mb) e um alto conteúdo de guanina e citosina (G-C), sendo, até o momento, o maior genoma viral já descrito. Dentre o conjunto de genes que compõem o grande genoma dos pandoravírus, uma grande parte é formada por ORFans, ou seja, são genes que não possuem homologia com nenhuma outra sequência depositada nos bancos de dados (PHILIPPE *et al.*, 2013).

Em seguida, outros vírus com partículas não icosaédricas também foram descritos. Os pithovírus foram isolados pela primeira vez a partir de amostras de *permafrost* da Sibéria, datadas de 30.000 anos atrás, e exibem partículas ovoides que podem chegar a 1,5 µm de comprimento (Figura 2 D) (LEGENDRE *et al.*, 2014). Já os cedratvírus apresentam partículas ovoides com tamanhos que podem alcançar 1,2 µm (Figura 2 E) e estão filogeneticamente relacionados com os pithovírus, o que levou à proposição de uma família para agrupá-los, chamada de Pithoviridae (ANDREANI *et al.*, 2016). Outro grupo de vírus gigantes, descoberto nas mesmas amostras de *permafrost* onde foram isolados os pithovírus, são os mollivírus. Esses vírus, descritos pela primeira vez em 2015, também apresentam suas peculiaridades, já que suas partículas de cerca de 500-600 nm de diâmetro são arredondadas (Figura 2 F). As relações evolutivas desses vírus com os outros NCLDV levaram à proposição de outra família hipotética para classificá-los, chamada de Molliviridae. (LEGENDRE *et al.*, 2015; CHRISTO-FOROUX *et al.*, 2020).

Vírus gigantes que parasitam outras amebas além daquelas do gênero *Acanthamoeba* também foram descritos na última década. É o caso dos Faustovírus, kaumoebavírus, orpheovírus e yasminevírus, que foram isolados a partir de amebas da espécie *Vermamoeba vermiformis*. Os faustovírus e kaumoebavírus estão filogeneticamente relacionados à família *Asfarviridae* e são vírus menores se comparados a maioria dos outros vírus de ameba, apresentando partículas com cerca de 200-250 nm (Figura 2 G). (RETENO *et al.*, 2015, BAJRAI *et al.*, 2016). Além disso, outro vírus gigante relacionado aos asfarvírus, mas que, no entanto, foi isolado em amebas do gênero *Acanthamoeba* é o pacmanvírus (ANDREANI *et al.*, 2017). Os orpheovírus são vírus também isolados em *Vermamoeba vermiformis*, mas diferentemente dos kaumoebavírus e Faustovírus eles possuem partícula (0,9 – 1,3 µm) e genoma maiores (~ 1.4 Mb) (Figura 2H), além de estarem filogeneticamente associados aos pithovírus (ANDREANI *et al.*, 2018). Já o yasminevírus, membro do grupo dos klosneuvírus

(família *Mimiviridae*), apresenta partículas de ~330 nm, que carregam um genoma extenso e complexo, composto por cerca de 2,1 Mb e codificando mais de mil proteínas, dentre as quais se encontra uma maquinaria quase completa para a tradução de proteínas (BAJRAI *et al.*, 2019).

No que diz respeito aos vírus gigantes brasileiros, o Grupo de Estudo e Prospecção de Vírus Gigantes (GEPVIG) da UFMG contribui, desde 2011, com a área. O primeiro vírus gigante a ser isolado no Brasil foi o sambavírus, membro da família *Mimiviridae*, que foi encontrado nas águas do Rio Negro, na Amazônia brasileira (CAMPOS *et al.*, 2014). Essa descoberta incentivou a busca por novos vírus gigantes no país, o que levou ao isolamento e caracterização de novas linhagens de mimivírus, marseillevírus, pandoravírus, cedratvírus e faustovírus, a partir de amostras ambientais de diferentes tipos. Isso evidencia a abundância desses vírus no Brasil (DORNAS *et al.*, 2015; ANDRADE *et al.*, 2018; ANDRADE *et al.*, 2019, RODRIGUES *et al.*, 2018a, BORGES *et al.*, 2019). Entretanto, a diversidade de vírus gigantes brasileiros ganhou maior destaque nos últimos anos, com a descoberta dos tupanvírus.

Os tupanvírus são membros da família *Mimiviridae* e foram isolados a partir de duas amostras diferentes, sendo uma delas de um lago salino do pantanal brasileiro e a outra de sedimentos que estavam a 3.000 metros de profundidade no Oceano Atlântico (ABRAHÃO *et al.*, 2018). Esses vírus apresentam uma partícula diferente do que era comumente observado para os mimivírus clássicos, já que possuem uma longa cauda associada ao seu capsídeo, o que faz com que algumas partículas possam alcançar até 2,3 µm de comprimento, tamanho que os torna os maiores vírus já descritos (Figura 3 A). O genoma dos tupanvírus é composto por 1,44 a 1,51 Mb, cujos genes codificam uma vasta maquinaria traducional, incluindo vários fatores de tradução e maturação de RNA, bem como 70 tipos de RNAt e 20 aminoacil-tRNA sintetases (aaRS). Esse conjunto de genes dos tupanvírus foi considerado o maior aparato traducional da virosfera conhecida até então (ABRAHÃO *et al.*, 2018). Outra característica diferencial dos tupanvírus é que eles parecem ter um espectro de hospedeiros mais expandido, sendo capazes de infectar tanto *Acanthamoeba* spp., quanto *Vermamoeba* spp. (ABRAHÃO *et al.*, 2018). Devido às análises filogenéticas e às características únicas dos tupanvírus, foi proposto que eles formassem um novo gênero dentro da família *Mimiviridae*, o gênero *Tupanvirus* (RODRIGUES *et al.*, 2018b).

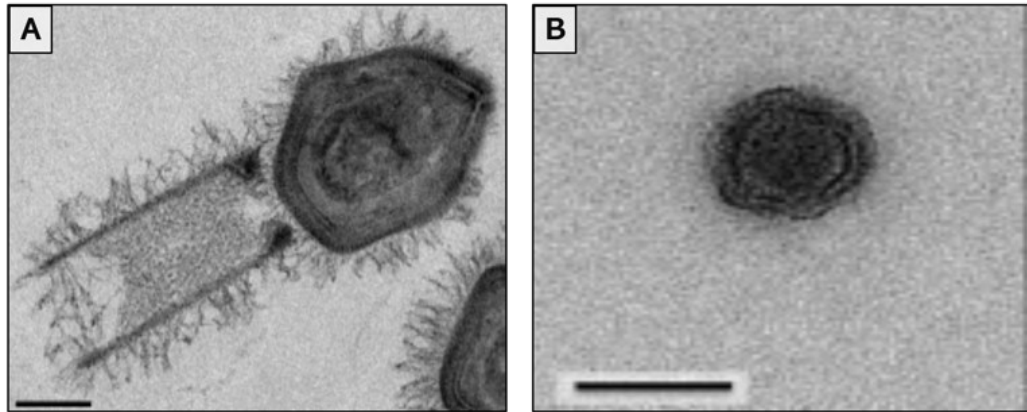


Figura 3 - Os dois vírus de ameba brasileiros que chamaram a atenção da comunidade científica. (A) Imagem de microscopia eletrônica de transmissão de uma partícula de tupanvírus, com sua cauda associada ao capsídeo. Barra de escala: 200 nm (B) Imagem de microscopia com contraste negativo, exibindo uma partícula isolada do yaravírus brasiliensis. Barra de escala: 100 nm. **Fonte:** Modificado de ABRAHÃO *et al.*, 2018 e BORATTO *et al.*, 2020.

Em contraste com a constante descoberta de vírus de ameba com grandes partículas e genomas está o isolamento do yaravírus brasiliensis, um vírus de 80 nm (Figura 3 B) que foi encontrado em amostras de água da Lagoa da Pampulha, em Belo Horizonte - MG (BORATTO *et al.*, 2020). A descoberta do yaravírus chamou bastante a atenção, pois o seu genoma de cerca de 44 kpb, é composto por 74 genes, dos quais 80% são ORFans. Apesar disso, esses vírus carregam em seu genoma algum aparato para tradução de proteínas, pois são capazes de codificar 6 diferentes RNAt (BORATTO *et al.*, 2020). Curiosamente, o yaravírus não apresenta determinadas características que unem os vírus de ameba ao grupo dos NCLDV, como grandes tamanhos de partícula e genoma e a presença de um conjunto de vários genes em comum. Dessa forma, ele poderia ser considerado o primeiro representante de vírus de *Acanthamoeba* spp. que não faz parte dos NCLDV, ou, em contrapartida, poderia até mesmo compor uma nova linhagem bastante divergente do grupo, o que o tornaria o menor NCLDV já descrito até então (BORATTO *et al.*, 2020).

A descoberta de tamanha diversidade de NCLDV nos últimos anos se tornou um grande desafio para a taxonomia desses organismos, que é realizada pelo Comitê Internacional em Taxonomia de Vírus (ICTV - *International Committee on Taxonomy of Viruses*). Como mencionado anteriormente, vários novos grupos virais foram sendo formados, levando à proposição de diversas famílias para incluir os novos isolados dentro do grupo dos NCLDV. Em meio às tentativas de classificá-los em um grupo taxonômico único, já foi proposta até a criação de uma ordem chamada de Megavirales, mas que não foi oficialmente reconhecida (COLSON *et al.*, 2012, COLSON *et al.*, 2013). Recentemente, o ICTV reclassificou muitas

famílias de vírus em táxons mais altos do que ordem, chegando a formar domínios (*realm*) e reinos (*kingdom*). Com base nisso, os vírus de DNAfd foram todos agrupados em um domínio chamado *Varidnaviria*, dentro do qual encontra-se o reino *Bamfordvirae*, que inclui os vírus cuja proteína do capsídeo apresenta a estrutura de *double-jelly-roll* (DJR), como é o caso dos NCLDV. O reino *Bamfordvirae* foi, então, dividido em dois filos, sendo que as famílias de NCLDV ficaram reunidas no filo *Nucleocytoviricota* (KOONIN *et al.*, 2020). A organização das famílias de NCLDV dentro do novo filo pode ser observada na Figura 4. No entanto, mesmo com a proposição dessa nova hierarquia taxonômica para os NCLDV, dentre os vírus de ameba, apenas as famílias *Mimiviridae* e *Marseilleviridae* são reconhecidas como oficiais pelo ICTV até então, embora a proposta de criação da família *Yaraviridae* tenha sido submetida e esteja sob análise pelo ICTV. Dessa forma, ainda é possível observar uma lacuna na taxonomia oficial dos diversos grupos de *Nucleocytoviricota* já descritos, que poderá ser mais bem elucidada com o acúmulo de novos estudos a respeito desses vírus.

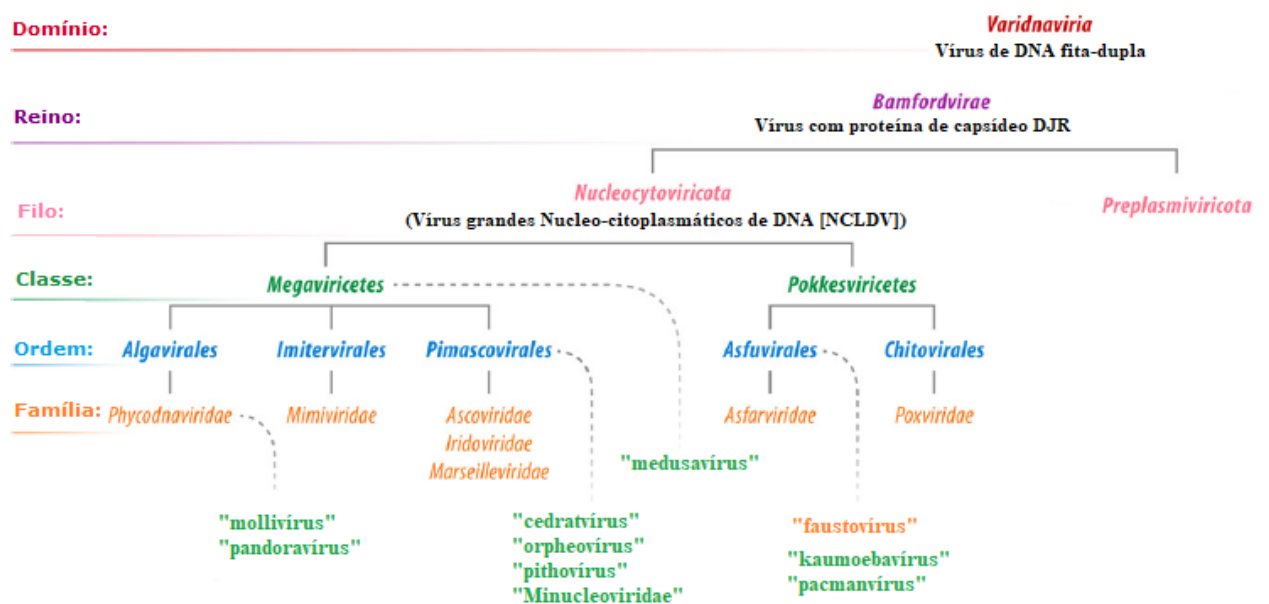


Figura 4 - Esquema representando a hierarquia taxonômica atual dos NCLDV (*Nucleocytoviricota*). Os grupos indicados entre aspas são aqueles que ainda não foram propostos ou reconhecidos oficialmente pelo ICTV. As linhas tracejadas representam os grupos cujas posições ainda são incertas. DJR: "double-jelly-roll". **Fonte:** Modificado de KOONIN *et al.*, 2020, p. 22.

1.3 Os mimivírus

1.3.1 A família *Mimiviridae*

Como indicado anteriormente, a família *Mimiviridae* é um táxon pertencente à ordem *Imitevirales*, à classe *Megaviricetes* e ao filo *Nucleocytoviricota*, que agrupa todos os NCLDV (KOONIN *et al.*, 2020). Oficialmente, a família abriga dois gêneros reconhecidos pelo ICTV: *Cafeteriavirus* e *Mimivirus* (ICTV ONLINE, 2020). O gênero *Cafeteriavirus* é composto pela espécie *Cafeteria roenbergensis virus* (CroV), que não é parasita de ameba e foi isolada em águas costeiras do Texas – EUA. O CroV infecta um protista heterotrófico flagelado marinho chamado *Cafeteria roenbergensis*. Esse vírus possui partículas de 300 nm e um genoma composto por cerca de 730 kpb, cujos genes codificam proteínas relacionadas à tradução e reparo de DNA, dentre outras (FISCHER *et al.*, 2010). Já o gênero *Mimivirus* é composto oficialmente por uma única espécie: *Acanthamoeba polyphaga mimivirus*, o APMV. Apesar disso, graças à intensificação dos estudos e aprimoramento das técnicas utilizadas para o isolamento de vírus gigantes, inúmeros isolados de mimivírus foram descritos, o que levou à descoberta de possíveis linhagens distintas desses vírus (PAGNIER *et al.*, 2013; KHALIL; ANDREANI; LA SCOLA, 2016). Ao analisar filogeneticamente esses diversos isolados de mimivírus conhecidos, é possível observar, inicialmente, a formação de três diferentes clados, chamados de linhagem A, B e C (Figura 5) (COLSON *et al.*, 2012). A linhagem A é representada pelo APMV e outros vírus mais parecidos, como os mamavírus e o isolado brasileiro sambavírus. Já a linhagem B é representada pelos moumouvírus e seus relacionados, enquanto a linhagem C é composta pelo megavírus chilensis e os outros vírus similares. (YOOSUF *et al.*, 2012; ARSLAN *et al.*, 2011).

Além dessas, outras linhagens vêm sendo identificadas (Figura 5), o que tem desencadeado várias propostas de atualização e organização da taxonomia desse grupo, como a, já mencionada, proposição de um novo gênero para abrigar os tupanvírus (RODRIGUES *et al.*, 2018b) e a possível divisão da família *Mimiviridae* em diferentes subfamílias (CLAVERIE; ABERGEL, 2018). A existência dessas possíveis subfamílias de *Mimiviridae* começou a ser descrita, principalmente, após o aumento da detecção de novas sequências de vírus por meio da metagenômica, método que tem contribuído bastante para a expansão dos membros da família. Nesse sentido, pode-se citar os klosneuvírus que são vírus inicialmente detectados por metagenômica e que apresentam um extenso aparato de genes envolvidos na tradução de proteínas, comparável ao dos tupanvírus (SCHULZ *et al.*, 2017). Outro exemplo é um conjunto de sequências de vírus de alga que foram inicialmente designadas como membros da família

Phycodnaviridae. Entretanto, análises posteriores mostraram que esses vírus possuíam mais genes homólogos com mimivírus do que com os phycodnavírus, sugerindo a sua inclusão em um novo clado de *Mimiviridae*, comumente chamado de “*extended-Mimiviridae*” (YUTIN *et al.*, 2013). Dessa forma, a grande diversidade da família *Mimiviridae* e o fato de esse grupo estar em constante expansão promovem a frequente atualização do conhecimento acerca da taxonomia, filogenia e evolução desse grupo, além de justificar a necessidade de novos estudos nessa área, que ajudarão a elucidar essas relações.

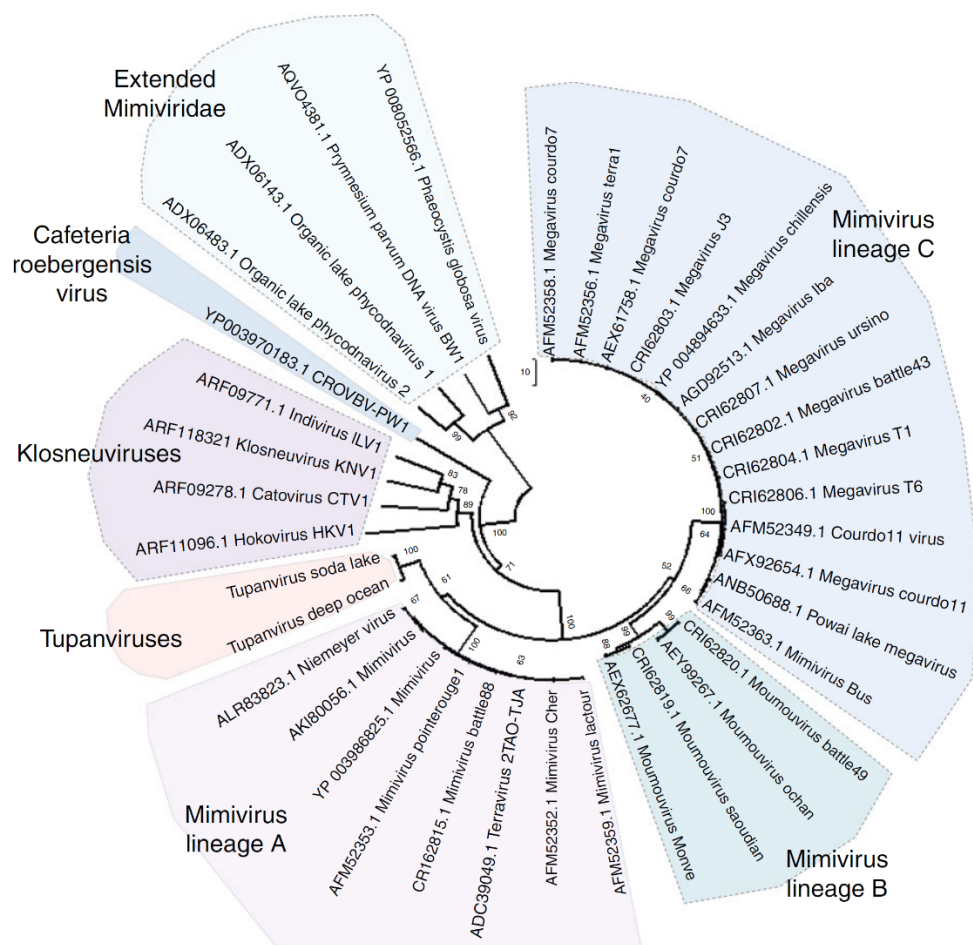


Figura 5 – Os diferentes clados da família *Mimiviridae*, representados em árvore filogenética, feita com base no gene que codifica a enzima DNA polimerase B. **Fonte:** Adaptado de ABRAHÃO *et al.*, 2018, p.4.

1.3.2 Estrutura da partícula do mimivírus

Na ocasião da descoberta do APMV, o primeiro isolado da família *Mimiviridae*, uma das características que mais chamou a atenção foi o aspecto de sua partícula viral, não só em questão de tamanho, mas também de complexidade. A partícula do APMV apresenta 750 nm

de tamanho e é composta por um capsídeo com simetria pseudoicosaédrica, do qual se projeta uma densa e compacta camada de longas fibrilas (Figura 6 A) (XIAO *et al.*, 2005, 2009). A estrutura do capsídeo é formada por várias camadas internas, sendo possível identificar uma membrana lipídica cercada por pelo menos três camadas de proteínas. Na parte interna na membrana lipídica existe um cerne (*core*) proteico, onde está localizado o genoma viral (Figura 6 B) (XIAO *et al.* 2009; ABRAHÃO *et al.*, 2014a).

A estrutura proteica do capsídeo dos mimivírus é formada, principalmente, pela proteína L425, que é a chamada proteína principal do capsídeo (MCP- *Major Capsid Protein*) (XIAO *et al.*, 2009). Em um dos vértices do capsídeo está presente uma estrutura que se projeta em formato de estrela-do-mar, chamada de *stargate* (Figura 6 C) (ZAUBERMAN *et al.*, 2008). A estrutura do *stargate* tem tamanho de 200-250 nm e é formada por proteínas diferentes da MCP. Ela é responsável pela liberação do conteúdo interno da partícula no citoplasma da célula hospedeira, funcionando como um portal para a liberação do genoma viral (ZAUBERMAN *et al.*, 2008). A região do capsídeo onde está posicionado o *stargate* é a única não recoberta pela camada de fibrilas (KLOSE *et al.*, 2010).

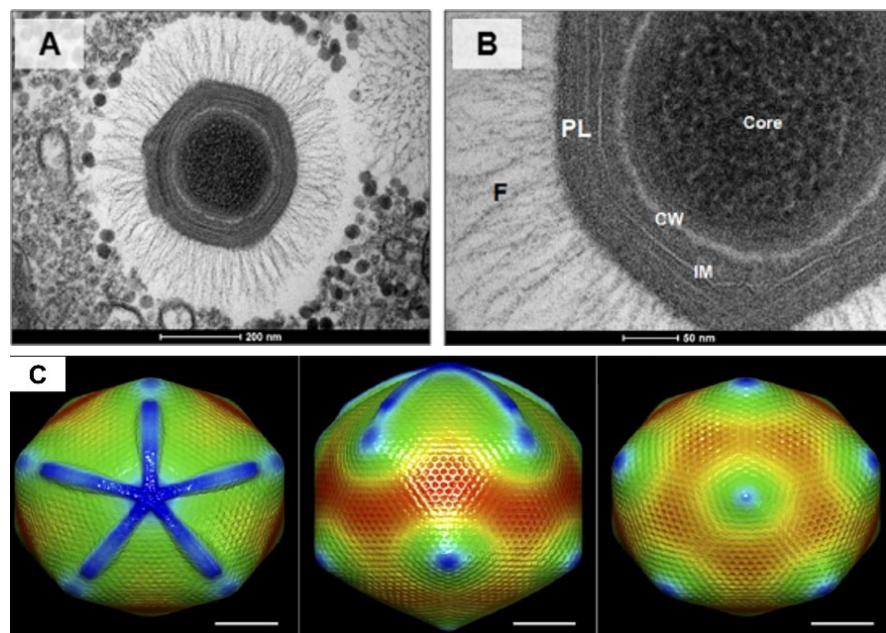


Figura 6 – Aspectos estruturais da partícula de um mimivírus. (A) Imagens de microscopia eletrônica de transmissão (MET), mostrando uma visão geral da partícula e exibindo a densa camada de fibrilas associada ao capsídeo viral **(B)** Imagem de MET da partícula, em maior escala, mostrando as fibrilas (F), as camadas de proteínas do capsídeo (PL), a membrana interna (IM) e a parede do *core* proteico (CW), que protege o genoma viral. **(C)** Reconstrução tridimensional da partícula do mimivírus, feita a partir de imagens de crio-microscopia eletrônica, que ilustra a estrutura *stargate* localizada em um dos vértices do capsídeo (em azul) e as demais regiões do capsídeo, sem o *stargate*. Barra de escala: 100 nm. **Fonte:** Modificado de ABRAHÃO *et al.*, 2014a, p.4 e XIAO *et al.*, 2009, p.964

Por fim, é importante ressaltar que, durante a sua morfogênese, as partículas infecciosas (vírions) dos mimivírus também são capazes de empacotar outras moléculas, além do genoma viral, necessárias para o início de um novo ciclo de multiplicação. Essas moléculas são os transcritos virais dos genes da DNA polimerase, da proteína principal do capsídeo (L425), de fatores de transcrição do tipo TFIIB, de três aminoacil-tRNA sintetases e de algumas ORFans (RAOULT *et al.*, 2004; SUZAN-MONTI; LA SCOLA; RAOULT, 2006). Além disso, análises de proteômica da partícula indicaram que os vírions de mimivírus carregam diversas proteínas em sua partícula, incluindo, além dos componentes estruturais já esperados, elementos da maquinaria de transcrição viral, como subunidades de RNA polimerase e fatores de transcrição, bem como diversas proteínas sem função conhecida (RENESTO *et al.*, 2006; CLAVERIE; ABERGEL; OGATA, 2009; SCHRAD *et al.*, 2020).

1.3.3 Ciclo de multiplicação do mimivírus

A primeira etapa da infecção por mimivírus se dá por meio da adesão inicial das partículas infecciosas com a superfície das células de ameba do gênero *Acanthamoeba*. Sugere-se que essa adesão possa estar relacionada com a presença da camada de fibrilas associada ao capsídeo viral (RODRIGUES *et al.*, 2015). A penetração do vírus acontece graças à fagocitose realizada pela célula hospedeira (SUZAN-MONTI *et al.*, 2007; ANDRADE *et al.*, 2017). Após a penetração, forma-se no citoplasma uma vesícula que engloba a partícula viral, chamada de fagossomo, que se funde com lisossomos celulares (ANDRADE *et al.*, 2017). O desnudamento da partícula e a liberação do conteúdo interno do capsídeo acontece após a abertura do *stargate* (ZAUBERMAN *et al.*, 2008). Algumas pesquisas indicam que a abertura completa do *stargate* é influenciada por uma combinação entre acidez e altas temperaturas (SCHRAD *et al.*, 2020).

Essa abertura do *stargate* promove uma fusão da membrana lipídica do fagossomo com a membrana interna da partícula viral. Isso permite a liberação da semente viral (cerne ou *core*) no citoplasma hospedeiro, liberando também o seu conteúdo interno, que é composto pelo DNA viral e por diversas proteínas importantes para os estágios iniciais da infecção (ANDRADE *et al.*, 2017; SCHRAD *et al.*, 2020). Nesse momento, não é possível encontrar partículas virais dentro da célula, uma vez que elas foram desnudadas. Essa etapa é chamada de fase de eclipse. Após a liberação da semente viral, se inicia uma reorganização do citoplasma, formando uma fábrica viral (FV) precoce, que é uma região elétron-densa onde ocorre a replicação e transcrição do DNA viral (ANDRADE *et al.*, 2017).

As proteínas virais são, então, sintetizadas no citoplasma e observa-se um considerável desenvolvimento da fábrica viral, que pode levá-la a ocupar 42% do volume celular em seu momento de maior maturação, estando dentre as maiores fábricas virais já descritas (SUZAN-MONTI *et al.*, 2007). As fábricas virais dos mimivírus são os centros de morfogênese das partículas virais da progênie, onde ocorre, principalmente, a síntese dos capsídeos, o preenchimento da partícula com o conteúdo que será carregado pelos vírions e a aquisição de fibrilas (KUZNETSOV *et al.*, 2013; ANDRADE *et al.*, 2017). A matéria-prima para a formação dos capsídeos vem da síntese proteica e do recrutamento de membranas celulares derivadas do retículo endoplasmático, que migram para as fábricas virais (MUTSAFI *et al.*, 2013).

Essas membranas se apresentam na forma de estruturas lamelares, presentes na periferia e no interior da fábrica. As estruturas lamelares, também chamadas de crescentes virais, vão adquirindo um formato mais angular e ganhando espessura, na medida em que migram do interior para a periferia da fábrica viral, o que leva à formação dos capsídeos de formato icosaédrico (ANDRADE *et al.*, 2017). Na periferia da fábrica viral os capsídeos recém-formados adquirem as suas fibrilas, ao mesmo tempo em que a partícula é preenchida com o genoma viral e as outras moléculas que serão carregadas no virion (ANDRADE *et al.*, 2017). Esse conteúdo interno do virion é adquirido por meio de uma região portal presente no lado oposto ao *stargate*, sendo esse último responsável apenas pela liberação do conteúdo interno e não pela sua aquisição (ZAUBERMAN *et al.*, 2008). A participação do núcleo celular em algum momento do ciclo dos mimivírus ainda é tema de debate (SUZAN-MONTI *et al.*, 2007; MUTSAFI *et al.*, 2010).

Ao fim do ciclo, quase todo o citoplasma está ocupado pela fábrica viral, estando a célula completamente preenchida por partículas recém-formadas. Por fim, as novas partículas virais são liberadas para uma nova infecção por meio da lise celular (SUZAN-MONTI *et al.*, 2007). O mecanismo de multiplicação dos mimivírus representa um importante alvo para diversos estudos desenvolvidos ao longo dos anos que sucederam a sua descoberta. Dessa forma, as etapas desse ciclo permanecem sendo exploradas e atualizações vão surgindo, na medida em que novas pesquisas são realizadas.

1.3.4 Genoma dos mimivírus

O genoma do APMV é composto por uma molécula linear de DNA dupla fita, que contém 1.181.549 pb, das quais 72% são Adenina-Timina (A-T) e 28% são Guanina-Citosina (G-C) (RAOULT *et al.*, 2004; LEGENDRE *et al.*, 2011). Diferentes análises genômicas mostraram

que esse vírus apresenta 1.018 genes, dentre os quais se encontram 979 sequências codificadoras de proteínas ou fases abertas de leitura (ORFs - *open reading frames*) e sequências preditas para codificar 6 RNAs transportadores, além de vários RNAs não codificantes. (LEGENDRE *et al.*, 2011). Os outros membros do gênero *Mimivírus*, correspondentes às linhagens A, B e C, apresentam características genômicas similares às do APMV (Tabela 1).

O repertório genômico dos mimivírus é composto por genes com homólogos tanto em outros vírus, quanto em bactérias, arqueias e eucariotos (ABERGEL; LEGENDRE; CLAVERIE, 2015). No que diz respeito aos genes com homólogos em outros vírus, um grupo importante é formado por sequências conservadas dentre os vírus que fazem parte do grupo dos NCLDV. Os produtos desses genes conservados estão associados à replicação de DNA, transcrição e morfogênese de partículas. Em meio a eles, é possível destacar a DNA polimerase B, comumente utilizada em análises filogenéticas do grupo, a primase/helicase D5, a ATPase A32L, associada ao empacotamento de conteúdo da partícula, além de um fator de transcrição e a proteína principal do capsídeo (IYER; ARAVIND; KOONIN, 2001; RAOULT *et al.*, 2004; YUTIN *et al.*, 2009).

Tabela 1 – Características gerais do genoma de alguns mimivírus das linhagens A, B e C.

Vírus	Linhagem	Tamanho do genoma (pb)	Conteúdo de G-C	ORFs preditas	Número de acesso (<i>GenBank</i>)
APMV	A	1.181.549	28%	979	NC_014649.1
Mamavírus	A	1.191.693	28%	1.023	JF801956.1
Samba vírus	A	1.181.380	27%	938	KF959826.2
Saudi mounovírus	B	1.030.056	25%	868	KY110734.1
Borely mounovírus	B	1.038.187	25%	943	MN175499
Megavírus chilensis	C	1.259.197	25%	1.120	NC_016072.1
Powai lake megavírus	C	1.208.707	25%	996	KU877344.1

APMV: *Acanthamoeba polyphaga* mimivírus. **Fonte:** Dados depositados no banco de dados do *National Center for Biotechnology Information* (NCBI) e extraídos de RAOULT *et al.*, 2004; LEGENDRE *et al.*, 2011; LA SCOLA *et al.*, 2008; CAMPOS *et al.*, 2014; BAJRAI *et al.*, 2016; SILVA *et al.*, 2020; ARSLAN *et al.* 2011 e CHATTERJEE *et al.*, 2019

Dentre os genes de mimivírus cujos homólogos pertencem a organismos não-virais, observa-se um grupo de sequências funcionalmente diverso, sendo que uma boa parte desses genes nunca havia sido descrita anteriormente em outros vírus (KOONIN, 2005). Esse conjunto de genes únicos de mimivírus foi classificado em quatro categorias genéricas funcionais: tradução de proteínas, enzimas de reparo do DNA, chaperonas e novas vias enzimáticas (RAOULT *et al.*, 2004). À primeira vista, o conjunto de genes de mimivírus que mais chamou

a atenção dos cientistas, foi o seu aparato de genes envolvidos na tradução de proteínas. O genoma dos mimivírus carrega homólogos para componentes importantes no processo, tais como aaRS, RNAs transportadores, fatores de tradução e uma enzima modificadora de RNAt (tRNA uracil 5-metiltransferase) (RAOULT *et al.*, 2004).

De maneira geral, os mimivírus da linhagem A, a qual inclui o APMV, apresentam 4 tipos de aaRS (Arginil-tRNA sintetase [ArgRS]; Cisteinil-tRNA sintetase [CysRS]; Metionil-tRNA sintetase [MetRS] e Tirosil-rRNA sintetase [TyrRS]), 5 tipos de fatores de tradução (IF4, IF4E, SUI1, eF-TU, eRF1) e 4 tipos de RNAt (leucina, histidina, cisteína e triptofano), sendo que o RNAt da leucina apresenta 3 duplicações, totalizando 6 sequências codificadoras de RNAt (RAOULT *et al.*; ABRAHÃO *et al.*, 2017). A quantidade desses genes varia bastante, tanto dentro da família *Mimiviridae*, quanto entre os membros do gênero *Mimivirus*, variando até mesmo entre isolados da mesma linhagem de mimivírus, já que alguns deles apresentam duplicações ou deleções de determinado gene. Por exemplo, o Niemeyer vírus (NYMV), isolado brasileiro de mimivírus da linhagem A, apresenta duplicações para três tipos de aaRS. Entretanto, ele possui apenas quatro sequências codificadoras de RNAt, ao invés de seis. Isso indica a presença de um aparato de tradução diferenciado em relação a outros membros da mesma linhagem, como o APMV (BORATTO *et al.*, 2015).

No que diz respeito às enzimas para reparo de DNA, essas são mais próximas a proteínas de eucariotos e bactérias (SUZAN-MONTI; LA SCOLA; RAOULT, 2006). O APMV apresenta 4 tipos de enzimas de reparo que nunca haviam sido descritas em vírus de DNAfd, sendo elas duas DNA glicosilases, que atuam localizando e excisando pirimidinas oxidadas (BANDARU *et al.*, 2007); uma endonuclease associada ao dano causado por radiação ultravioleta; uma proteína da família MutS, envolvida no reparo de incompatibilidade (*mismatch*) de DNA e um tipo de metiltransferase (6-O-metilguanina-DNA-metiltransferase), descrita pela primeira vez em vírus (RAOULT *et al.*, 2004). Além disso, os mimivírus também codificam topoisomerases de 3 tipos (II A, I A, IB), sendo que a topoisomerase do tipo IA do APMV foi a primeira descrita na virosfera. Ademais, a presença simultânea desses três tipos de topoisomerases é rara em microrganismos (RAOULT *et al.*, 2004).

Quanto às chaperonas, enzimas relacionadas ao enovelamento de proteínas, os mimivírus apresentam homólogos de diversos tipos, como com a família HSP70 (DnaK) de chaperonas e com proteínas com domínio DnaJ. Adicionalmente, o APMV foi o primeiro vírus a apresentar sequências homólogas a um tipo de proteína do choque térmico (*heat-shock*) da bactéria

Escherichia coli e a uma ciclofilina (peptidilprolil isomerase). Essa última parece ter características únicas e conservadas dentro do gênero *Mimivirus*, o que levou a proposição de uma nova família de ciclofilinas. (RAOULT *et al.*, 2004; THAI *et al.*, 2008; BARIK, 2018).

A última categoria de proteínas únicas dos mimivírus, proposta por Raoult e colaboradores (2004), é a das vias enzimáticas. Dentre essas vias metabólicas codificadas pelo genoma do mimivírus está o metabolismo de aminoácidos, lipídeos e açúcares. Em meio às enzimas do APMV, envolvidas no metabolismo de aminoácidos, estão a asparagina sintase e a glutamina sintase (RAOULT *et al.*, 2004). Já no caso do metabolismo de lipídeos, os mimivírus codificam uma enzima homóloga à lanosterol 14-alfademetilase, que foi a primeira proteína da superfamília citocromo P450 descrita em um vírus (LAMB *et al.*, 2009). Além disso, os mimivírus também codificam homólogos para uma 7-desidrocolesterol redutase e uma colinesterase. Quanto ao metabolismo de açúcares, vale destacar a presença da enzima GlmU, que catalisa a síntese de um metabólito central para a formação de todos os tipos de polissacarídeos. Além de outras enzimas envolvidas na síntese e modificação de carboidratos, o APMV codifica homólogos para 6 tipos de glicosiltransferases, que são enzimas que participam da glicosilação de proteínas após a tradução. Muitas das proteínas dos mimivírus são altamente glicosiladas, incluindo a MCP e as glicoproteínas que compõem a camada de fibrilas da partícula (RAOULT *et al.*, 2004; KLOSE *et al.*, 2010).

No contexto da maquinaria de glicosilação dos mimivírus, a enzima procolágeno-lisina 2-oxoglutarato 5-dioxigenase, cujo homólogo é codificado pela ORF L230 do APMV, foi inicialmente predita como uma proteína que poderia modificar colágeno após a tradução (RAOULT *et al.*, 2004). Estudos posteriores concluíram que ela é uma proteína bifuncional, com função de lisil hidroxilase e glicosiltransferase de colágeno e que, de fato, atua na modificação pós traducional do colágeno. A L230 foi descrita, então, como a primeira enzima de glicosilação de colágeno viral, sendo um novo tipo dessa proteína na natureza, além de ser a primeira hidroxilisina não-animal descrita (LUTHER *et al.*, 2011). Os colágenos formam uma família de proteínas muito abundante nos animais, pois além de sua função estrutural nos tecidos, eles atuam também como reguladores de muitos processos celulares (MYLLYHARJU *et al.*, 2004).

A estrutura dessas proteínas do tipo colágeno é constituída por três cadeias (cadeias-alfa), formadas por repetições de 3 resíduos de aminoácidos, Glicina-X-Y. A posição X é, preferencialmente, composta por prolina e a posição Y, geralmente, composta por 4-

hidroxiprolina. No entanto, as posições X e Y do triplex de aminoácidos podem ser compostas por qualquer outro aminoácido diferente de glicina. Essas três cadeias se enrolam uma sobre as outras no sentido horário, formando uma tripla hélice de colágeno. (MYLLYHARJU *et al.*, 2004; NELSON; COX, 2019). Além dos animais, proteínas tipo colágeno já foram descritas em fungos, bactérias e arqueias (LUTHER *et al.*, 2011). O genoma do APMV possui oito genes para proteínas contendo repetições que formam a estrutura de tripla hélice de colágeno e antes da descrição dos mimivírus, proteínas tipo colágeno virais só eram observadas em bacteriófagos (RAOULT *et al.*, 2004, LUTHER *et al.*, 2011). Apesar de alguns trabalhos indicarem que proteínas com repetição de tripla hélice de colágeno possam estar associadas à superfície das partículas de mimivírus, mais estudos são necessários para elucidar o papel dessas proteínas para a biologia dos mimivírus (RAOULT *et al.*, 2004; SHAH *et al.*, 2014).

É interessante destacar que o conteúdo proteico dos mimivírus também inclui vários outros grupos de proteínas contendo domínios com repetições específicas, além daquelas tipo colágeno. Por exemplo, um dos motivos mais frequentes dentre as proteínas de mimivírus está uma repetição de anquirina (ANK), encontrada em mais de 30 ORFs do APMV e predita para estar associada com interações entre proteínas (RAOULT *et al.*, 2004). Outros tipos de repetições distribuídas pelo proteoma de mimivírus são: a repetição Kelch; a repetição rica em leucina (LRR- *leucine-rich repeat*), repetição do tetratricopeptídeo (TPR); repetição MORN (*membrane occupation, and recognition nexus*); repetição de fenilalanina-asparagina-isoleucina-prolina (FNIP/IP22); repetição triptofano-ácido aspártico (WD40) e a repetição Sel 1. A presença de proteínas contendo essas repetições já foi positivamente correlacionada com o tamanho do genoma dos vírus gigantes, ou seja, quanto maior o genoma, maior a quantidade de proteínas contendo essas estruturas (SHUKLA *et al.*, 2018)

Apesar do extenso repertório de genes dos mimivírus aqui descrito, responsável por codificar várias proteínas com funções e origem já propostas, a maior parte do genoma desses vírus é composta por ORFans. No caso do APMV, por exemplo, mais da metade de seus genes preditos não apresentava homólogos nos bancos de dados na época de sua descrição. Por causa disso, novas tentativas de determinar funções para essas ORFans, utilizando novas metodologias, têm sido feitas (SAINI *et al.*, 2007; HAKIM *et al.*, 2012; SOBHY *et al.*, 2015). Isso demonstra que, aos poucos e com novos estudos, esses genes completamente desconhecidos podem sim ter as suas funções e origens desvendadas. Por fim, a presença de um conjunto diverso de genes associados a funções celulares tão importantes, bem como seu vasto conteúdo de ORFans, indicam a singularidade dos mimivírus na virosfera.

1.4 Os virófagos

1.4.1 A descoberta dos virófagos

A descoberta dos mimivírus abriu portas para a descrição de uma diversidade de novas linhagens e de novos grupos de vírus gigantes. O primeiro vírus do gênero *Mimivirus* descrito após o isolamento do APMV, foi o *Acanthamoeba castellanii* mamavírus (ACMV ou mamavírus). O mamavírus foi isolado a partir da inoculação de amostras de água de uma torre de refrigeração de Paris – França, em células de *Acanthamoeba* sp. (LA SCOLA *et al.*, 2008). No entanto, as imagens de microscopia eletrônica da ameba infectada indicaram que, além das partículas do mamavírus, estavam presentes dentro da célula partículas virais menores, até então desconhecidas (LA SCOLA *et al.*, 2008). Essas pequenas partículas, com cerca de 50 nm, chamaram a atenção dos pesquisadores devido a sua peculiar relação com o mamavírus e com a ameba hospedeira.

Os pesquisadores observaram que esses vírus menores se multiplicavam dentro da fábrica viral dos mamavírus e que não eram capazes de se multiplicar sozinhos na ameba, ou seja, sem uma coinfeção com mamavírus ou com APMV. Além disso, os estudos também evidenciaram que a presença dessas partículas menores acabava afetando o ciclo de multiplicação do mamavírus, já que parecia aumentar significativamente a formação de partículas de mamavírus defeituosas, o que causava um impacto para a produtividade do vírus gigante e para a sobrevivência das amebas (LA SCOLA *et al.*, 2008). Para definir a natureza dessas pequenas partículas recém-descobertas, os pesquisadores franceses propuseram o termo virófago, fazendo referência aos bacteriófagos, que são os vírus que infectam bactérias. Dessa forma, o primeiro virófago descrito foi chamado de sputnik (LA SCOLA *et al.*, 2008).

Por definição, os virófagos são vírus de DNAfd, com partículas de 50 a 74 nm, que requerem, além da célula hospedeira, a presença de um vírus gigante para estabelecer o seu próprio ciclo de multiplicação. (KRUPOVIC; KUHN; FISCHER, 2015). Dessa forma, eles foram considerados parasitas virais ou vírus que infectam outros vírus, já que, ao invés de utilizarem a maquinaria de transcrição da célula hospedeira, eles exploram a fábrica viral e a maquinaria do vírus gigante para garantir a expressão de seus genes. Isso pode, em certos casos, causar impactos negativos para formação da progênie do vírus gigante (DESNUES; RAOULT, 2012; MOUGARI *et al.*, 2019c).

A criação do termo virófago, proposta por La Scola e colaboradores (2008), foi um dos fatores responsáveis por desencadear um debate acerca da natureza desse novo tipo de vírus. A

principal pauta do debate era se os virófagos deveriam mesmo ser considerados como uma nova entidade viral, ou se seriam apenas novos isolados dos, já conhecidos, vírus satélites. Até então, os vírus satélites eram definidos pelo ICTV como agentes subvirais, compostos por moléculas de ácidos nucleicos, e que dependem da coinfeção de célula hospedeira com um vírus auxiliador para se multiplicar (DESNUES; RAOULT, 2010). Por um lado, pesquisadores defendiam que partículas como o sputnik não apresentavam características únicas o suficiente para serem classificadas como uma entidade diferente de vírus satélites (KRUPOVIC; CVIRKAITE-KRUPOVIC, 2011). Por outro lado, outros pesquisadores defendiam que os virófagos eram muito complexos e que apresentavam características únicas para serem definidos como vírus satélites, ou seja, simples agentes subvirais (FISCHER, 2011). De forma a resolver esse impasse, ambos os lados do debate acabaram concluindo que a classificação proposta pelo ICTV para os vírus satélites e, portanto, para os virófagos deveria ser revista. Assim, novas propostas de classificação para os virófagos e vírus satélites começaram a ser discutidas.

Inicialmente, foi sugerida uma classificação baseada no sistema dicotômico proposto por Raoult e Forterre (2008), que definia os organismos codificadores de capsídeos (CEOs) e os organismos codificadores de ribossomos (REOs). Dessa forma, seria criado um grupo chamado de “vírus de vírus”, dentro dos CEOs. Esse grupo incluiria tanto os virófagos, quanto os vírus satélites conhecidos (DESNUES; RAOULT, 2012). Entretanto, foi a proposta de um outro esquema de reclassificação taxonômica para os virófagos e vírus satélites, que recebeu o aceite oficial do ICTV. A ideia dessa proposta era que ambos fossem considerados como vírus e que nenhum fosse considerado agente subviral, já que são capazes de codificar o seu próprio capsídeo. Assim, foi proposto que os virófagos fossem classificados em uma nova família viral chamada de *Lavidaviridae*, enquanto os vírus satélites seriam separados em duas famílias distintas, *Virgaviridae* e *Sarothroviridae* (KRUPOVIC; KUHN; FISCHER, 2015).

Com a recente reorganização taxonômica realizada pelo ICTV, essas famílias também foram agrupadas em táxons superiores, assim como os *Mimiviridae* e outros vírus gigantes. As famílias de vírus satélite foram agrupadas no domínio *Riboviria*, que reúne os vírus com genoma de RNA. Já a família *Lavidaviridae*, composta pelos virófagos, foi incluída no domínio *Varidnaviria*, junto aos outros vírus de DNAdf. Como a proteína principal do capsídeo dos virófagos apresenta a estrutura DJR, o grupo também foi incluído no reino *Bamfordvirae*, formando o filo *Preplasmaviricota*, ao lado dos *Nucleocytoviricota*. Por fim, a família

Lavidaviridae está incluída dentro da classe *Maveriviricetes* e da ordem *Priklausovirales* (Figura 7) (KOONIN *et al.*, 2020)

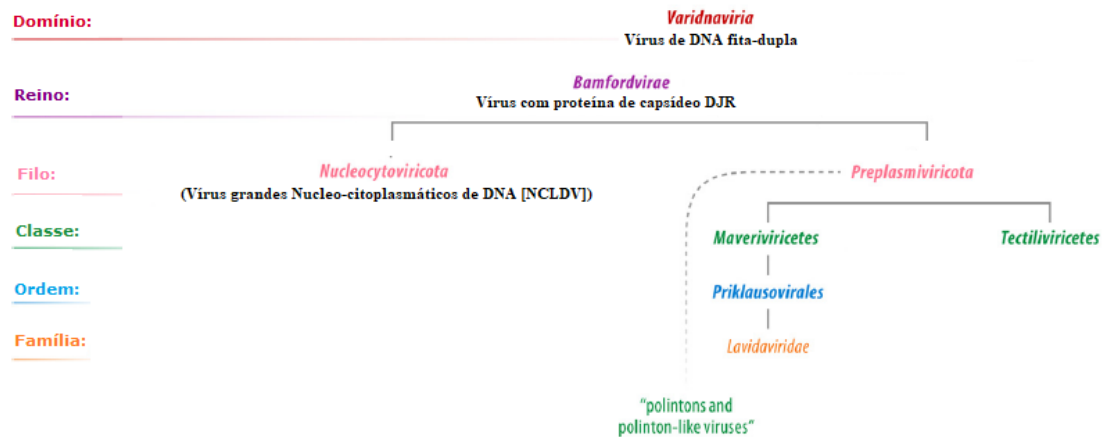


Figura 7 - Esquema indicando a taxonomia atual dos virófagos. O grupo indicado entre aspas ainda não foi proposto ou reconhecido oficialmente pelo ICTV. A linha tracejada representa posição ainda incerta para o grupo. DJR: “double-jelly-roll”. **Fonte:** Modificado de KOONIN *et al.*, 2020, p. 22.

1.4.2 A família *Lavidaviridae*

Os membros da família *Lavidaviridae*, criada para abrigar os virófagos, são caracterizados por compartilhar um conjunto de proteínas e domínios conservados, codificados pelos *core genes* do grupo. Dentre as proteínas ou domínios codificados pelos *core genes*, destacam-se a proteína principal do capsídeo (MCP), uma proteína secundária do capsídeo (mCP - *minor capsid protein*), uma ATPase da família FtsK-HerA associada ao empacotamento de DNA e uma protease. A conservação e a monofilia dessas proteínas sugerem que os virófagos possam compartilhar uma ancestralidade comum, sustentando a criação de uma nova família para agrupá-los (KRUPOVIC; KUHN; FISCHER, 2015; YUTIN; RAOULT; KOONIN, 2013).

A proposta de formação da família *Lavidaviridae* foi acompanhada pelo estabelecimento de dois critérios importantes para definir quais vírus podem fazer parte da família ou não. O primeiro critério preconiza que os vírus devem possuir pelo menos alguns dos *core genes* associados à morfogênese da partícula que são conservados no grupo. Assim, esses genes podem ser usados para análises filogenéticas. O segundo critério propõe que o vírus candidato a fazer parte da família dependa ou esteja associado a um vírus do grupo dos NCLDVs (KRUPOVIC; KUHN; FISCHER, 2015).

A descoberta do virófago sputnik, naturalmente, abriu portas para a posterior descrição de outros virófagos. O segundo virófago isolado, o mavírus, foi obtido a partir de águas costeiras do Texas, nos Estados Unidos, associado ao Cafeteria roenbergensis vírus (CroV) (FISCHER; SUTTLE, 2011). Algum tempo depois, dois novos isolados de virófago sputnik, chamados de sputnik 2 e sputnik 3, foram descobertos na França, a partir de amostras de fluido de lentes de contato e de amostras de solo, respectivamente (DESNUES *et al.*, 2012; GAIA *et al.*, 2013a). O virófago sputnik 2 foi isolado em associação a um novo mimivírus, o lentillevírus, enquanto o virófago sputnik 3 foi descoberto por meio de um procedimento que visou o isolamento de novos virófagos utilizando como plataforma o cocultivo de células de *Acanthamoeba polyphaga* infectadas com um mimivírus (DESNUES *et al.*, 2012; GAIA *et al.*, 2013a). Além disso, três virófagos de uma outra linhagem que infecta mimivírus foram descritos e chamados de zamilon, zamilon 2 e zamilon vitis (GAIA *et al.*, 2014; BEKLIZ *et al.*, 2015; JEUDY *et al.*, 2020).

O zamilon foi o primeiro virófago descoberto em associação a um mimivírus da linhagem C e, por meio dos estudos realizados, foi observado que ele não era capaz de se multiplicar em coinfeção com mimivírus da linhagem A, o que foi uma descoberta intrigante e que se tornou tema de investigação (BOUGHALMI *et al.*, 2013; GAIA *et al.*, 2014). Pesquisas desenvolvidas, posteriormente, propuseram a existência de um mecanismo de defesa contra virófagos nos mimivírus da linhagem A (LEVASSEUR *et al.*, 2016). Foram detectadas pequenas repetições, derivadas do genoma do Zamilon, na sequência do gene R349 de todos os mimivírus da linhagem A (LEVASSEUR *et al.*, 2016). Esse gene, quando silenciado, reestabelece a susceptibilidade dos mimivírus de linhagem A à infecção pelo virófago zamilon (MOUGARI *et al.*, 2019b). Além disso, o estudo da região do DNA de mimivírus próxima às sequências de Zamilon detectadas identificou duas proteínas com função de nuclease e helicase, que estão relacionadas às famílias Cas3 e Cas4 de enzimas associadas ao sistema de defesa das bactérias contra bacteriófagos (CRISPR-Cas) (LEVASSEUR *et al.*, 2016; DOU *et al.*, 2018). Desse modo, foi proposto que os mimivírus da linhagem A utilizam um mecanismo de defesa viral similar ao CRISPR-Cas, que foi chamado por eles de sistema MIMIVIRE (Elemento de resistência a virófagos dos mimivirus - *MIMIVirus Virophage Resistance Element*). Apesar disso, o mecanismo de defesa viral utilizando o sistema MIMIVIRE ainda é tema de debate. (CLAVERIE; ABERGEL, 2016).

É importante ressaltar que o desenvolvimento de vários estudos com abordagem metagenômica, cujo objetivo é detectar sequências de novos virófagos em diferentes ambientes,

tem sido bastante significativo ao contribuir com a expansão da família *Lavidaviridae*. O primeiro virófago descrito por meio da metagenômica foi o “Organic Lake Virophage” (OLV). Ele foi detectado em um lago orgânico hipersalino da Antártida e propõe-se que esse virófago parasite vírus membros da família *Phycodnaviridae* (YAU *et al.*, 2011). Sete sequências completas de virófagos também foram detectadas utilizando bancos de dados de metagenômica do lago Yellowstone, nos Estados Unidos. Esses virófagos foram chamados de “Yellowstone Lake virophages” (YSLVs) e devido às diferenças genotípicas entre as suas sequências, eles foram numerados de 1 a 7 (YSLV1 – YSLV7) (ZHOU *et al.*, 2013; ZHOU *et al.*, 2014). Junto à detecção dos primeiros YSLVs, mas a partir de amostra de um lago da Antártida, também foi obtida a sequência quase completa dos “Ace Lake Mavírus” (ALM), que compartilham 13 genes homólogos com os mavírus (ZHOU *et al.*, 2013). Posteriormente, novas sequências de virófagos continuaram sendo detectadas em diversos trabalhos, como é o caso do “Phaeocystis globosa virus virophage” (PgVV) e dos achados chineses “Dishui Lake Virophage” (DLV) e “Qinghai Lake Virophage” (QLV) (SANTINI *et al.*, 2013; GONG *et al.*, 2016; OH; YOO; LIU, 2016).

Além de genomas completos, algumas sequências incompletas ou genes específicos de virófagos também vêm sendo descritas e detectadas por metagenômica. Essa abordagem metagenômica tem permitido a detecção de virófagos em ambientes muito diversos e, até mesmo, extremos. Esses ambientes incluem, além daqueles já citados, o solo hiperárido de deserto da Antártida, buracos cheios de água (*cryoconites*) presentes em geleiras e até mesmo o rúmen de ovelhas (ZABLOCKI *et al.*, 2014; BELLAS; ANESIO; BARKER, 2015; YUTIN; KAPITONOV; KOONIN, 2015). Em conjunto, esses estudos evidenciam que os virófagos podem estar presentes em diversos ambientes ao redor do mundo e que novos estudos visando o isolamento das partículas nesses ambientes podem ser bem-sucedidos.

No território brasileiro, o primeiro virófago, chamado de virófago Rio Negro, foi descoberto em associação com o sambavírus, a partir de amostras de água do Rio Negro, na Amazônia (CAMPOS *et al.*, 2014). O segundo virófago brasileiro, chamado de virófago Guarani, foi isolado a partir de amostras de água da Lagoa da Pampulha, em Belo Horizonte, MG (MOUGARI *et al.*, 2019a). Ambos os virófagos brasileiros possuem genoma de aproximadamente 18 mil pares de bases, que contém entre 20 e 22 ORFs. As análises de genoma desses virófagos mostram que eles são genotipicamente similares ao virófago sputnik. O conjunto dos virófagos já detectados ou isolados e algumas de suas características estão listados na tabela 2.

Tabela 2– Lista de virófagos descritos e algumas de suas principais características.

Virófago	Vírus associado	Célula hospedeira	Tipo de amostra (país)	Tamanho do genoma (número de ORFs)	Referência
Sputnik*	<i>Acanthamoeba castellanii</i> mamavírus	<i>Acanthamoeba</i> sp.	Água de torre de refrigeração (França)	18.343 pb (21)	LA SCOLA <i>et al.</i> , 2008
Organic Lake virophage (OLV)	Organic Lake phycodnavírus (<i>extended Mimiviridae</i>)?	Algas verdes marinhas?	Lago orgânico hipersalino (Antártica).	26.421 pb (26)	YAU <i>et al.</i> , 2011
Mavírus*	<i>Cafeteria roenbergensis</i> vírus (CroV)	<i>Cafeteria roenbergensis</i>	Águas costeiras do Texas (EUA)	19.063 pb (20)	FISCHER; SUTTLE, 2011
Sputnik 2*	Lentillevírus	<i>Acanthamoeba</i> sp.	Fluido de lentes de contato (França)	18.338 pb (21)	DESNUES <i>et al.</i> , 2012
YellowStone Lake Virophage 1 (YSLV1)	<i>Phycodnaviridae</i> ou <i>Mimiviridae</i> ?	Microalgas?	Lago Yellowstone (EUA)	27.849 pb (26)	ZHOU <i>et al.</i> , 2013
YellowStone Lake Virophage 2 (YSLV2)	<i>Phycodnaviridae</i> ou <i>Mimiviridae</i> ?	Microalgas?	Lago Yellowstone (EUA)	23.184 pb (21)	ZHOU <i>et al.</i> , 2013
YellowStone Lake Virophage 3 (YSLV3)	<i>Phycodnaviridae</i> ou <i>Mimiviridae</i> ?	Microalgas?	Lago Yellowstone (EUA)	27.050 pb (23)	ZHOU <i>et al.</i> , 2013
YellowStone Lake Virophage 4 (YSLV4)	<i>Phycodnaviridae</i> ou <i>Mimiviridae</i> ?	Microalgas?	Lago Yellowstone (EUA)	28.396 (34)	ZHOU <i>et al.</i> , 2013
Ace Lake Mavirus (ALM)	<i>Mimiviridae</i> ?	Protistas fagotróficos?	Lago Ace (Antártica)	17.767 (22)	ZHOU <i>et al.</i> , 2013
Sputnik 3*	Linhagens A, B e C do gênero <i>Mimivirus</i>	<i>Acanthamoeba polyphaga</i>	Amostras de solo (França)	18.338 pb (21)	GAIA <i>et al.</i> , 2013a
<i>Phaeocystis globosa</i> vírus virophage (PgVV)	<i>PgV-16T</i> (<i>extended Mimiviridae</i>)?	<i>Phaeocystis globosa</i> ?	Águas costeiras (Holanda)	19.527 (16)	SANTINI <i>et al.</i> , 2013
Zamilon*	Mont1 vírus (mimivírus linhagem C)	<i>Acanthamoeba polyphaga</i>	Amostras de solo (Tunísia)	17.276 pb (20)	BOUGHALMi <i>et al.</i> , 2013; GAIA <i>et al.</i> , 2014
Rio Negro *	Samba vírus	<i>Acanthamoeba castellanii</i>	Amostras de água doce do Rio Negro (Brasil).	18.145 pb (20)	CAMPOS <i>et al.</i> , 2014; BORGES <i>et al.</i> , 2018
YellowStone Lake Virophage 5 (YSLV5)	Desconhecido	Desconhecido	Lago Yellowstone (EUA)	29.767pb (32)	ZHOU <i>et al.</i> , 2014

(Continua...)

Virófago	Vírus associado	Célula hospedeira	Tipo de amostra (país)	Tamanho do genoma (número de ORFs)	Referência
YellowStone Lake Virophage 6 (YSLV6)	Desconhecido	Desconhecido	Lago Yellowstone (EUA)	24.837 pb (29)	ZHOU <i>et al.</i> , 2014
YellowStone Lake Virophage 7 (YSLV7)	Desconhecido	Desconhecido	Lago Yellowstone (EUA)	23.193 pb (26)	ZHOU <i>et al.</i> , 2014
Bigelowiella natans virophages (BnVP)**	esconhecido	<i>Bigelowiella natans?</i>	-	0,1– 33,3 kpb	BLANC <i>et al.</i> , 2015
Rumem Virophages (RVPs)**	<i>Mimiviridae?</i>	Protistas habitantes do rúmen	Rúmen de ovelha (EUA)	-	YUTIN; KAPITONOV; KOONIN, 2015
Cryoconite virophage**	NCLDV não classificados?	Amebas?	Buracos em geleiras – <i>cryoconites</i> (Groelândia)	12.595 pb	BELLAS <i>et al.</i> , 2015
Zamilon 2**	<i>Mimiviridae?</i>	<i>Acanthamoeba</i> spp.	Biorreator (EUA)	6.716 pb (15)	BEKLIZ <i>et al.</i> , 2015
Dishui Lake Virophage (DSLIV)	Desconhecido	Microalgas?	Lago Dishui (China)	28.788 pb (28)	GONG <i>et al.</i> , 2016
Qinghai Lake Virophage (QLV)	Desconhecido	Desconhecido	Lago Qinghai (China)	23.379 pb (25)	OH; YOO; LIU, 2016
Lake Mendota and Trout Bog Lake virophages	NCLDV?	Desconhecido	Lagos Mendota e Trout Bog (EUA)	13,8 – 25,8 kpb (13-25)	ROUX <i>et al.</i> , 2017
Platanovirus saccamoebae virophage*	Platanovirus Saccamoebae - KSL5(<i>Mimiviridae</i>)	<i>Saccamoeba</i> sp	Plantas do gênero <i>Platanus</i> (Alemanha)	-	MICHEL <i>et al.</i> , 2018
CpV- associated virophages (CpV-PLVs)	<i>Chrysochromulina parva</i> virus (CpV-BQ2)	<i>Chrysochromulina parva</i>	Lago Ontário (América do Norte)	21.750- 22.879 pb (19-23)	STOUGH <i>et al.</i> , 2019
Guarani*	<i>Mimiviridae</i>	<i>Acanthamoeba</i> spp.	Lagoa da Pampulha (Brasil)	18.967 pb (22)	MOUGARI <i>et al.</i> , 2019a
Zamilon vitis*	Megavírus vitis	<i>Acanthamoeba</i> spp.	Solo de vinhedo (França)	17.454 pb (20)	JEUDY <i>et al.</i> , 2019
NYC virophages**	Desconhecido	Desconhecido	Esgoto (EUA)	-	GULINO <i>et al.</i> , 2020

Notas: PgV-16T: *Phaeocystis globosa* virus -16T; pb: pares de bases; kpb: quilo pares de bases; NCLDV: vírus grandes núcleo-citoplasmáticos de DNA; CpV: *Chrysochromulina parva* vírus; NYC: New York City

* Virófagos que foram isolados, utilizando métodos de cultivo. Os demais foram detectados por meio de estudos de genômica ou metagenômica.

** Detecção de genoma parcial ou de fragmentos e genes específicos (geralmente, proteína principal do capsídeo)
? Indica que foi proposto, mas não confirmado por métodos de cultivo.

Fonte: Informações adaptadas de BEKLIZ; COLSON; LA SCOLA, 2016; FISCHER, 2020; MOUGARI *et al.*, 2019c e dos trabalhos citados na coluna Referência.

Apesar de tantos virófagos já identificados ao redor do mundo, o ICTV reconhece apenas dois gêneros para a família *Lavidaviridae*. O gênero *Mavirus* é composto oficialmente pela espécie *Cafeteriavirus-dependent mavirus* e foi proposto para agrupar os virófagos tipo mavirus. Já o segundo gênero, chamado *Sputnikvirus*, é formado oficialmente por duas espécies: *Mimivirus-dependent virus Sputnik* e *Mimivirus-dependent virus Zamilon*, propostas para agrupar virófagos tipo sputnik e tipo zamilon, respectivamente (KRUPOVIC; KUHN; FISCHER, 2015). A classificação oficial ainda não considera os virófagos detectados por metagenômica, como o OLV e as linhagens de YSLV, no entanto, foi proposto que eles compusessem gêneros diferentes na família (KRUPOVIC; KUHN; FISCHER, 2015; DUPONCHEL; FISCHER, 2019). Isso demonstra que a expansão da família dos virófagos, graças às novas descobertas, torna necessários os constantes estudos sobre a filogenia, evolução e classificação desses organismos.

1.5 Os virófagos sputnik

1.5.1 Estrutura da partícula

Estudos de crio-microscopia eletrônica das partículas do primeiro virófago sputnik isolado, mostraram que eles apresentam partículas não envelopadas, de aproximadamente 74 nm de diâmetro e que têm a simetria icosaédrica (Figura 8) (SUN *et al.*, 2009). O capsídeo é composto por unidades proteicas, chamadas de capsômeros, que são compostas por duas proteínas diferentes, as proteínas principal e secundária do capsídeo (MCP e mCP, respectivamente). A maioria dos capsômeros são formados por trímeros da MCP, que apresenta estrutura de DJR, e se organizam em 260 unidades pseudo-hexaméricas (hexons). Já os capsômeros compostos pela mCP formam 12 unidades pentaméricas (pentons). Na parte central dos capsômeros pentaméricos existem cavidades que podem servir como portal para entrada e saída do material genético do virófago, assim como acontece em bacteriófagos (SUN *et al.*, 2009; ZHANG *et al.*, 2012).

As análises da estrutura do virófago sputnik permitiram observar que no centro de cada capsômero pseudo-hexamérico existem protuberâncias, com cabeça triangular, chamadas de fibras do tipo cogumelo, que têm 5.5 nm de comprimento e cobrem a superfície da partícula

(SUN *et al.*, 2009). Hipotetiza-se que as funções dessas fibras possam estar relacionadas à adesão do sputnik às partículas do vírus gigante hospedeiro, garantindo a penetração na célula (DESNUES; BOYER; RAOULT, 2012). Esses estudos propuseram também a existência de uma membrana lipídica dentro do capsídeo do sputnik, já que foi detectada uma camada mais densa dentro do capsídeo e que esses virófagos contêm entre 12% e 24 % do seu peso composto por lipídeos (SUN *et al.*, 2009). No entanto, estudos posteriores feitos com a mesma metodologia, mas utilizando uma melhor resolução de imagem, mostram que a partícula do sputnik na verdade não apresenta uma membrana lipídica interna ao capsídeo (ZHANG *et al.*, 2012).

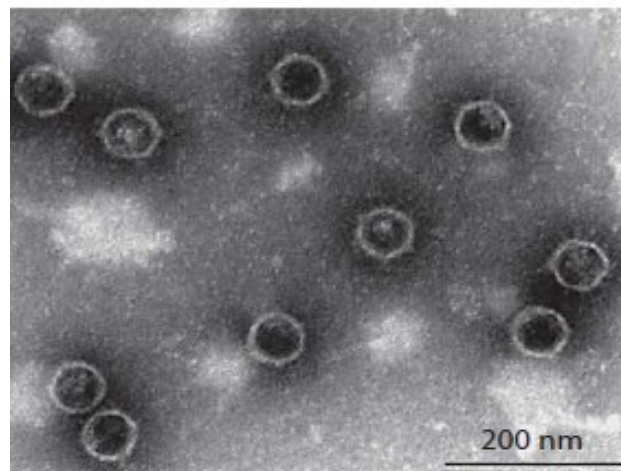


Figura 8 - Imagem de microscopia eletrônica de transmissão de partículas do virófago sputnik obtidas utilizando contraste negativo. **Fonte:** Modificado de DESNUES; RAOULT, 2010, p.296

Além disso, os estudos de Zhang e colaboradores (2012) mostraram a presença de três camadas paralelas que estão presentes dentro do capsídeo do virófago sputnik, cada uma com aproximadamente 2,5 nm de espessura, sendo a camada mais externa mais densa do que as mais internas. Esses pesquisadores propuseram que essas três camadas representam o DNA de dupla fita ordenado ou semi-ordenado e que o fato de a camada mais externa ser mais densa pode ser explicado pela associação íntima do DNA à estrutura do capsídeo (ZHANG *et al.*, 2012). Outros estudos revelaram ainda que é possível que no conteúdo interno da partícula do sputnik seja encontrado um conjunto de RNAs virais empacotados e prontos para uso durante o ciclo de multiplicação (DESNUES; RAOULT, 2010). Um resumo geral da estrutura da partícula de sputnik está representado na Figura 9.

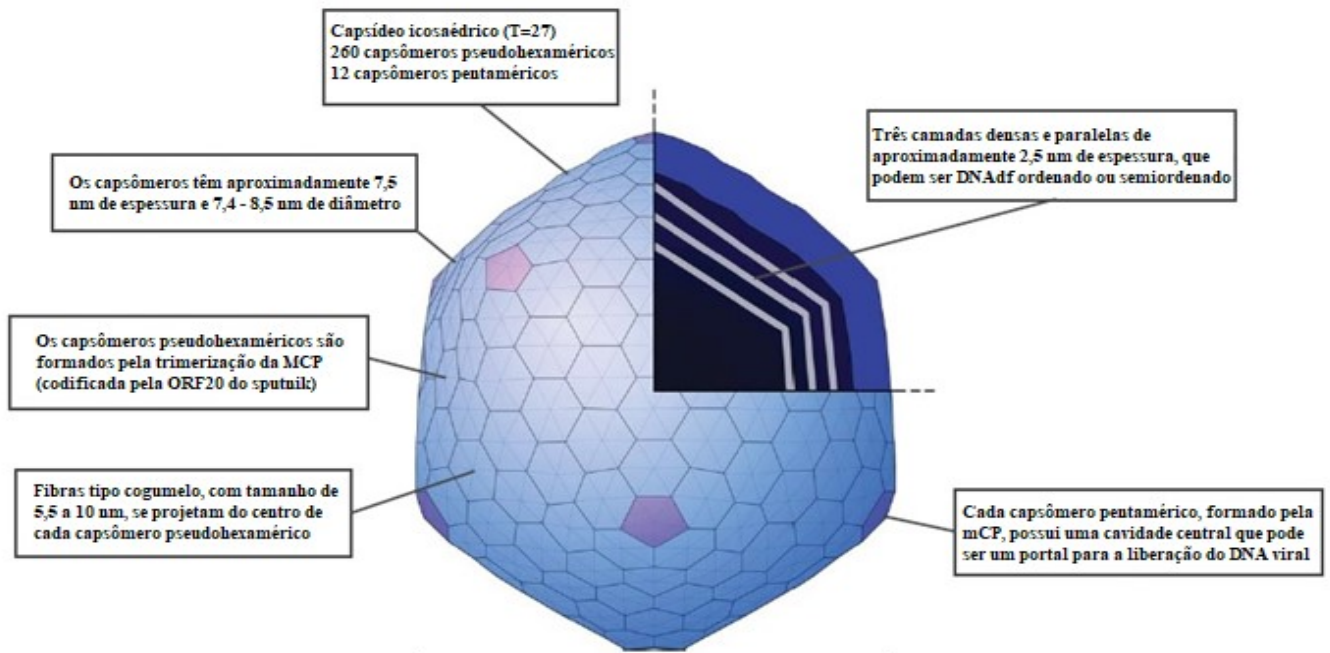


Figura 9 - Esquema que resume a estrutura geral de partícula do viróforo sputnik. T: número de triangulação; MCP: proteína principal do capsídeo; mCP: proteína secundária do capsídeo. **Fonte:** Modificado de GAIA *et al.*, 2013b, p4

1.5.2 Ciclo de multiplicação dos viróforos sputnik

A estratégia de multiplicação dos viróforos é caracterizada por ser dependente da presença de um vírus gigante para se estabelecer. Dessa forma, o ciclo de multiplicação do viróforo sputnik está diretamente associado ao dos mimivírus (LA SCOLA *et al.*, 2008). A infecção se inicia com a penetração da partícula do viróforo sputnik na ameba. Até o momento, a principal hipótese para explicar a ocorrência dessa etapa é a de que o viróforo sputnik penetre na célula por fagocitose, por meio de uma copenetração com o vírus gigante. Essa copenetração ocorreria devido a uma interação entre as fibras do capsídeo do viróforo sputnik com a superfície de fibrilas do mimivírus, unindo as duas partículas. Essa hipótese é sustentada pela observação de partículas do sputnik associadas às fibrilas do APMV e pela incapacidade de multiplicação do viróforo sputnik quando em coinfeção com vírus sem fibrilas, como o mimivírus M4 e os marseillévirus. (DESNUES; RAOULT, 2010; GAIA *et al.*, 2013b; BOYER *et al.*, 2011). Entretanto, mais estudos são necessários para que essa etapa inicial seja mais bem esclarecida, considerando outras formas de penetração para o sputnik, já que os viróforos mavírus são capazes de penetrar na célula utilizando um mecanismo independente do vírus gigante, a endocitose mediada por clatrina (FISCHER; SUTTLE, 2011).

Logo após a infecção, os virófagos já podem ser observados no citoplasma da ameba, localizados dentro de vesículas endocíticas. (DESNUES; RAOULT, 2010; BEKLIZ; COLSON; LA SCOLA, 2016). A liberação do genoma do virófago para o citoplasma celular é, provavelmente, realizada por meio das cavidades centrais presentes nos capsômeros pentaméricos que compõem o capsídeo do virófago sputnik, local por onde o material genético também seria encapsidado (SUN *et al.*, 2009; ZHANG *et al.*, 2012). Esse processo de liberação pode ser desencadeado em resposta a estresses no ambiente celular, como por exemplo, uma queda do pH, de maneira similar ao que acontece com outros vírus icosaédricos, como os adenovírus (ZHANG *et al.*, 2012).

Com o genoma liberado, a replicação do DNA, a transcrição e a tradução das proteínas se iniciam no período entre 2 e 4 horas pós-infecção, quando se observa a fase de eclipse (BEKLIZ; COLSON; LA SCOLA, 2016). Sugere-se que a replicação do DNA dos virófagos seja catalisada por DNA polimerases e helicases codificadas pelo próprio virófago, enquanto a expressão gênica seja iniciada utilizando fatores de transcrição codificados pelo vírus gigante, durante fases finais da infecção (DUPONCHEL; FISCHER, 2019). Após a fase de eclipse, a morfogênese das novas partículas de virófago ocorre em um dos polos da fábrica viral do mimivírus já formada. (DESNUES; RAOULT, 2010; GAIA *et al.*, 2013b; BEKLIZ; COLSON; LA SCOLA, 2016).

A formação de nova progênie dos mimivírus parece ocorrer após a morfogênese dos virófagos, em outro polo da fábrica viral. Decorridas 16 horas do início da infecção já é possível detectar a progênie de ambos os vírus preenchendo a célula da ameba, sendo que o virófago pode ser observado no citoplasma ou dentro de vesículas celulares. Por fim, a liberação do virófago sputnik é feita junto com a do mimivírus, por meio da lise celular. Entretanto, assim como para a penetração, outros mecanismos de liberação, como a exocitose, ainda precisam ser mais bem estudados (DESNUES; RAOULT, 2010; BEKLIZ; COLSON; LA SCOLA, 2016). O conjunto das etapas propostas para o ciclo de multiplicação dos virófagos está ilustrado na Figura 10.

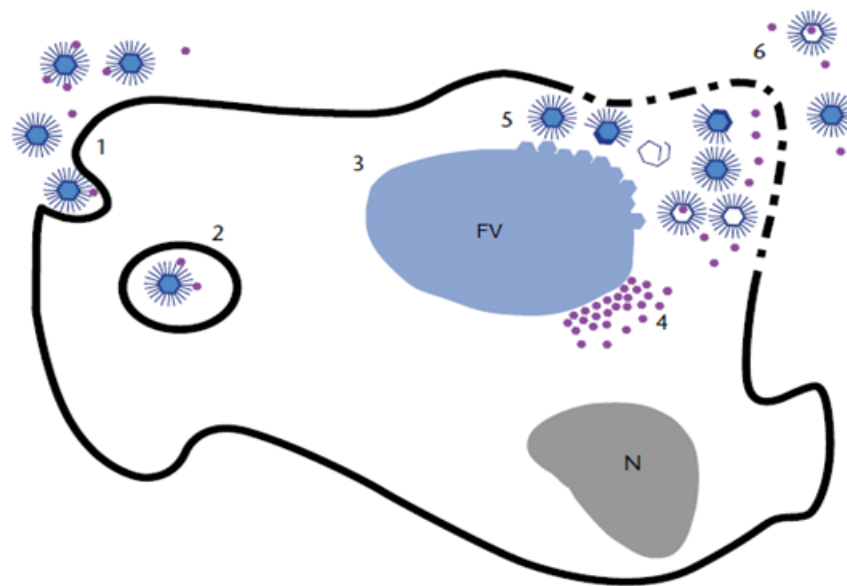


Figura 10 - Esquema representativo dos diferentes passos da coinfeção de uma célula de ameba pelo virófago sputnik e por um mimivírus. (1) Copenetração de mimivírus e virófago sputnik na ameba, por meio da fagocitose. (2) Ambos os vírus são internalizados em vacúolos celulares. (3) Surgimento da FV do mimivírus (4) Produção de partículas do virófago em um dos pólos da FV do mimivírus. (5) Posterior produção de novas partículas de mimivírus, sendo que algumas delas podem ser defeituosas. (6) Liberação da progênie de ambos os vírus, por meio da lise da célula. N= núcleo. FV: fábrica viral. **Fonte:** Modificado de GAIA *et al.*, 2013b, p.6

1.5.3 Genoma dos virófagos sputnik

Os genomas dos três isolados de virófago sputnik (1, 2 e 3) apresentam 99% de similaridade um com o outro e cada um é composto por uma molécula de DNA de dupla fita circular (GAIA *et al.*, 2013b). O virófago sputnik 1 tem seu genoma formado por 18.343 pb, sendo que 72% dessas bases são adenina e timina (A+T) (LA SCOLA *et al.*, 2008). Tanto o genoma do sputnik 2 quanto o do sputnik 3 são compostos por 18.338 pb (DESNUES *et al.*, 2012; GAIA *et al.*, 2013a). Assim como o primeiro sputnik, os outros dois isolados também apresentam um conteúdo de A+T maior do que 70% e essa característica também está presente nos mimivírus (GAIA *et al.*, 2013b). Os virófagos tipo sputnik isolados no Brasil, o rio negro e o guarani, possuem genoma com 18.145 pb e 18.967 pb, respectivamente. Ambos apresentam conteúdos genéticos altamente similares aos dos virófagos sputnik, ilustrando a presença dessas entidades em ambientes diversos (BORGES *et al.*, 2018; MOUGARI *et al.*, 2019a). Além disso, foi predito que os virófagos sputnik 1, 2 e 3 são capazes de codificar 21 ORFs, enquanto os virófagos rio negro e guarani apresentam 20 e 22 sequências codificadoras, respectivamente. (LA SCOLA *et al.*, 2008; DESNUES *et al.*, 2012; GAIA *et al.*, 2013a; MOUGARI *et al.*, 2019a).

Os produtos desses genes nos virófagos sputnik podem ser divididos em dois diferentes grupos, o das proteínas conservadas e o das proteínas não conservadas entre todos os virófagos. As proteínas conservadas estão associadas à morfogênese da partícula dos virófagos, e a sua conservação e monofilia sugerem que os virófagos evoluíram a partir de um ancestral comum. Dentre essas proteínas, se inclui a MCP, a mCP, uma ATPase e uma protease (YUTIN; RAOULT; KOONIN, 2013). Tanto a MCP quanto a mCP compõem a estrutura básica do capsídeo, sendo que a MCP é codificada pela ORF 20 nos virófagos sputnik 1, 2, 3 e no rio negro (LA SCOLA *et al.*, 2008; DESNUES *et al.*, 2012; GAIA *et al.*, 2013a). Já no virófago Guarani, duas ORFs (ORF1 e ORF 22) codificam a proteína principal do capsídeo (MOUGARI *et al.*, 2019a). Diferentemente da MCP, a mCP apresenta estrutura de *single-jelly-roll*. Estudos da estrutura da partícula do virófago sputnik mostraram que a mCP que forma as estruturas pentaméricas do capsídeo (também chamada de proteína *penton*) é produto de uma combinação das ORFs 18 e 19 de sputnik 1, 2 e 3, o que indica que, na verdade, esses três isolados apresentariam 20 ORFs ao invés de 21 (ZHANG *et al.*, 2012). Em todos os isolados de sputnik, uma outra proteína estrutural, também considerada como proteína secundária do capsídeo, é codificada por uma sequência diferente da ORF 18/19. O par de proteínas de capsídeo MCP-mCP é bastante conservado, e está presente em todos os virófagos, o que evidencia a sua importância para esses organismos (FISCHER, 2020). Além disso, não existem homólogos detectáveis para essas proteínas em outros organismos, além dos virófagos, o que também é indicativo de uma ancestralidade em comum (YUTIN; RAOULT; KOONIN, 2013).

Durante a morfogênese da partícula na fábrica viral, as outras duas proteínas conservadas citadas anteriormente têm papel muito importante. A cisteína protease é responsável por clivar a MCP na porção C-terminal, garantindo a maturação do capsídeo (BORN *et al.*, 2018). Já a ATPase da superfamília FtsK-HerA foi proposta para estar envolvida no empacotamento do DNA para dentro da partícula (LA SCOLA *et al.*, 2008). Outras duas proteínas podem ser consideradas conservadas dentro dos virófagos, apesar de não estarem presentes em todos eles e de suas relações filogenéticas ainda não serem bem esclarecidas. É o caso de uma proteína com domínio de helicase da superfamília 3 (S3H), o qual está acoplado a um outro domínio amino-terminal com atividade predita de primase-polimerase (TVpol). Essa proteína, possivelmente, cumpre o papel de enzima primária de replicação de DNA nos virófagos sputnik. (FISCHER, 2020). Além disso, duas ORFs dos virófagos sputnik codificam uma proteína contendo fita de zinco (ZnR – *Zinc-ribbon*) que também é, frequentemente, codificada

pelos virófagos e foi, inicialmente, predita para atuar na regulação da transcrição (YUTIN; RAOULT; KOONIN, 2013; LA SCOLA *et al.*, 2008)

Quanto às proteínas não conservadas entre todos os virófagos, que estão presentes no genoma dos isolados de virófago sputnik, pode-se citar duas ORFs que codificam proteínas parálogas contendo tripla hélice de colágeno (ORFs 6 e 7), cujos homólogos são proteínas do mesmo tipo, presentes em mimivírus (LA SCOLA *et al.*, 2008). Propõe-se que, nos virófagos, a presença desse tipo de proteína, que contém repetições e que são homólogas às proteínas de seu vírus hospedeiro, possa estar relacionada com a interação entre virófago e vírus gigante. (YUTIN; RAOULT; KOONIN, 2013; LA SCOLA *et al.*, 2008). Além disso, funções e características de três outras ORFs também foram preditas, sendo elas uma integrase, uma transposase e uma proteína de membrana. A integrase da família tirosina (tyr) recombinase pode estar relacionada com uma possível integração da sequência do virófago ao genoma do vírus gigante ao qual está associado, enquanto a transposase da família IS3 é uma enzima que catalisa o movimento de elementos genéticos móveis, como transposons, no genoma (LA SCOLA *et al.*, 2008). É interessante destacar que, assim como nos mimivírus, uma parte dos genes dos virófagos sputnik também não apresentaram homólogos nos bancos de dados analisados até então, o que caracteriza esses genes como ORFans. (LA SCOLA *et al.*, 2008)

Em conjunto, os genes dos virófagos sputnik apresentam ligação evolutiva com vírus que infectam os três domínios da vida (LA SCOLA *et al.*, 2008). Isso sugere que os virófagos façam constantes transferências horizontais de genes com os vírus gigantes e com a ameba hospedeira. Além disso, os virófagos também podem trocar genes com os elementos genéticos móveis, que são sequências que migram entre um genoma e outro, utilizando enzimas específicas. Esse fato é considerado um evento chave para a evolução dos virófagos. (KOONIN; KRUPOVIC, 2017). Dessa forma, o sucesso evolutivo dos virófagos pode estar intimamente relacionado com a sua constante transferência gênica com outros organismos e com os elementos genéticos móveis, bem como com a sua capacidade de se integrar seu genoma no DNA de seus hospedeiros (DUPONCHEL; FISCHER, 2019).

1.6 Os elementos genéticos móveis

Os elementos genéticos móveis (EGMs) são quaisquer segmentos de DNA capazes de se mover dentro de um mesmo genoma ou entre genomas diferentes, por meio da codificação de enzimas e outras proteínas que fazem a mediação deste movimento (SIEFERT, 2009; FROST *et al.*, 2005). O conjunto de todos os elementos genéticos móveis de um genoma é chamado de mobiloma (SIEFERT, 2009). De acordo com Siefert (2009), os EGMs podem ser divididos em quatro grandes grupos. O primeiro grupo reúne os transposons (ou elementos transponíveis). O segundo engloba os plasmídeos, enquanto o terceiro inclui os elementos derivados de vírus, como os profagos originados de infecções de bactérias por bacteriófagos. Já o quarto grupo é formado pelos parasitas moleculares *self-splicing*, como os íntrons e inteínas.

A primeira descrição de um EGM foi feita pela cientista Bárbara McClintock, após a realização de experimentos que envolviam cruzamentos entre grãos de milho, na década de 1940. Ela observou que os padrões de pigmentação dos grãos de milho eram modificados após os cruzamentos e propôs, assim, a existência de elementos, chamados por ela de elementos controladores, que se moviam no cromossomo do milho e alteravam a expressão de alguns genes (MCCLINTOCK, 1950). Essa descoberta foi feita em um período no qual se acreditava que os genes mantinham sua posição estável e linear dentro dos cromossomos, o que levou à resistência da comunidade científica para aceitar a mudança desse paradigma da genética. Entretanto, a posterior descoberta de EGMs em outros organismos, além do milho, contribuiu para que o trabalho de McClintock fosse reconhecido, rendendo à pesquisadora o Prêmio Nobel de Fisiologia ou Medicina de 1983, trinta e três anos após sua principal publicação sobre o assunto. Desde então, com o avanço das tecnologias de biologia molecular, foi possível melhor estudar o papel dos EGMs para os genomas nos quais eles habitam. Atualmente, sabe-se que os EGMs são abundantes e ubíquos, pois estão amplamente distribuídos nos genomas de diferentes espécies, que abrange desde plantas e animais até o vírus. (RAVINDRAN, 2012)

Por algum tempo, os elementos móveis foram considerados simples parasitas moleculares, pois se acreditava que eles não forneciam qualquer benefício para seu hospedeiro, sendo apenas um tipo de DNA “lixo” ou “egoísta”, que avoluma as sequências genômicas e que visa garantir a sua própria replicação, sem causar alterações fenotípicas (SCHRADER; SCHMITZ, 2018; KIDWELL; LISCH, 2001). Entretanto, com o avanço das técnicas moleculares e genômicas, a percepção sobre a função dos EGMs foi se modificando. Mesmo que ainda sejam considerados parasitas moleculares, as mudanças genômicas geradas por eles

representam uma importante contribuição para a evolução dos genomas que os hospedam (MILLER; CAPY, 2004).

O papel dos EGMs na evolução pode ser exemplificado pelo desenvolvimento de mecanismos de defesa contra o EGM pelo genoma hospedeiro. Esses mecanismos incluem estratégias de metilação das sequências móveis, além da expressão de moléculas, como RNAs, que são capazes de suprimir a atividade do EGM (LEVIN; MORAN, 2011). Por outro lado, os elementos móveis também desenvolveram mecanismos de direcionamento bastante específicos, que os permitem se integrar ao genoma em locais onde eles conseguem escapar dos mecanismos de defesa desenvolvidos contra eles (LEVIN; MORAN, 2011). Essa corrida armamentista é importante do ponto vista evolutivo, pois desencadeia inovações nas formas de vida celular, o que evidencia o papel positivo dos EGMs na evolução (KOONIN, 2016).

Além disso, os EGMs são importantes agentes da transmissão gênica horizontal (TGH) e da geração de variação genética. No caso das bactérias, por exemplo, transposons, profagos e plasmídeos podem ser transmitidos por meio dos processos de transformação, transdução e conjugação. Uma importante aplicação desses eventos nos procariotos é a transferência de genes de resistência aos antimicrobianos de uma bactéria para a outra, principalmente utilizando plasmídeos conjugativos (FROST *et al.*, 2005). Dessa forma, a presença e atividade dos elementos genéticos móveis em diferentes genomas, podem estar associadas à diversificação genética e ao surgimento de novas espécies ou variantes adaptativas, que estarão sujeitas à seleção natural, tornando os EGMs importantes impulsionadores da evolução genômica dos organismos (SCHRADER; SCHMITZ, 2018).

1.6.1 Transposons

Os transposons, também chamados de elementos transponíveis (ETs), são o principal o tipo de elemento genético móvel presente nos genomas eucarióticos. Por meio do processo de transposição e replicação, os ETs podem afetar a dinâmica genômica de uma célula de diversas maneiras. Ao atuar como promotores, silenciadores ou como sítios de modificações epigenéticas, por exemplo, os transposons podem alterar a expressão gênica da célula. Além disso, as diversas cópias de ETs presentes em um único genoma, podem funcionar como *hotspots* para eventos de recombinação e afetar o tamanho da sequência (SIEFERT, 2009; WICKER *et al.*, 2007).

Com base em diferentes critérios relacionados ao processo de transposição, os ETs podem ser classificados hierarquicamente em classes, subclasses, ordens, superfamílias, famílias e subfamílias, conforme proposto por Wicker e colaboradores em 2007. O nível hierárquico mais alto são as classes, que são divididas com base no uso ou não de uma molécula de RNA como intermediária para a transposição. Já os níveis mais baixos da hierarquia se dividem com base em diferentes critérios, como os mecanismos de transposição e inserção, tipos de enzima utilizada, estratégia de replicação, conservação da sequência de DNA e filogenia. Assim, os transposons podem ser divididos em duas grandes classes (WICKER *et al.*, 2007).

A classe I é composta pelos chamados retrotransposons, que utilizam uma molécula de RNA como intermediária para a transposição. O mecanismo de transposição e replicação dos retrotransposons é dependente da enzima transcriptase reversa. Essa enzima utiliza a molécula de RNA, transcrita a partir do genoma do retrotransposon, para sintetizar uma nova molécula de DNA, produzindo novas cópias desse elemento transponível. Assim, a cada ciclo de replicação, o elemento transponível é duplicado e a nova cópia é inserida no genoma-alvo, mantendo a sequência original no genoma doador. Esse tipo de mecanismo é chamado de “copiar e colar” (WICKER *et al.*, 2007; SCHRADER; SCHMITZ, 2018). Os retrotransposons podem ser divididos em vários subgrupos, os quais apresentam características distintas entre si, como a presença ou não de repetições terminais longas (LTR, *long terminal repeats*) e de enzimas como integrase ou recombinase, relacionadas com seus mecanismos de integração no DNA alvo (WICKER *et al.*, 2007).

A classe II agrupa transposons que não utilizam uma molécula de RNA para intermediar a sua transposição, os quais também podem ser chamados de DNA transposons. Essa classe de transposons pode ser subdividida em duas subclasses, de acordo com o tipo de transposição realizada (WICKER *et al.*, 2007). A primeira dessas subclasses é composta pelos DNA transposons que se movem por meio do mecanismo de “cortar e colar”. Para tanto, eles utilizam a enzima transposase, que reconhece as repetições terminais invertidas (TIRs, *terminal inverted repeats*) presentes em suas extremidades. Assim, a função da transposase é cortar as duas fitas do DNA doador em cada uma das extremidades do transposon e em seguida inserir a sequência excisada no DNA alvo (WICKER *et al.*, 2007; SCHRADER; SCHMITZ, 2018). Ao contrário do mecanismo de “copiar e colar”, o mecanismo de “cortar e colar” dessa subclasse de DNA transposons não deixa a sequência do transposon original inserida no DNA doador. Nesse caso, os DNA transposons se aproveitam da replicação do genoma hospedeiro para realizar a sua

transposição. Assim, eles podem se mover de uma região já duplicada para uma que ainda não passou por esse processo, aumentando, dessa forma, o seu número de cópias (WICKER *et al.*, 2007).

Diferentemente da primeira subclasse de DNA transposons, a segunda subclasse agrupa aqueles transposons cuja transposição envolve a clivagem de apenas uma das fitas do DNA, mantendo a sequência do transposon inserida no DNA doador em um mecanismo de “copiar e colar”. Dois grupos de DNA transposons distintos fazem parte dessa subclasse: os helintrons e os Maverick/Polintons (WICKER *et al.*, 2007). Os helintrons foram os primeiros DNA transposons que não utilizam o mecanismo de “cortar e colar” descritos. O seu mecanismo de transposição envolve a replicação do tipo círculo rolante e uma de suas características é a capacidade de codificar uma proteína com atividade de DNA helicase, bem como a ausência das repetições terminais invertidas (KAPITONOV; JURKA, 2001).

Já os transposons do tipo *Maverick*, também chamados frequentemente de polintons, são grandes sequências móveis, cuja molécula de DNA pode apresentar até 25 kpb, sendo considerados os maiores genomas dentre os transposons conhecidos e até mesmo chamados de elementos transponíveis gigantes (KRUPOVIC; KOONIN, 2016; PRITHAM; PUTLIWALA; FESCHOTTE, 2006; FESCHOTTE; PRITHAM, 2007). Os polintons são capazes de codificar vários genes, cujos produtos apresentam diferentes funções. Dentre eles, pode-se destacar o gene para uma integrase tipo retrovírus (RVE) e para uma DNA polimerase B *protein-primed*, que são conservados dentre esses elementos. Por codificarem proteínas-chave para a sua própria replicação, os polintons são considerados transposons autorreplicativos. Além disso, os polintons também codificam sequências homólogas a proteínas virais e estão conectados evolutivamente com os vírus grandes de DNA e com os virófagos (KRUPOVIC; KOONIN, 2016). Um esquema que resume os níveis de classificação de transposons aqui referidos pode ser observado na figura 11.

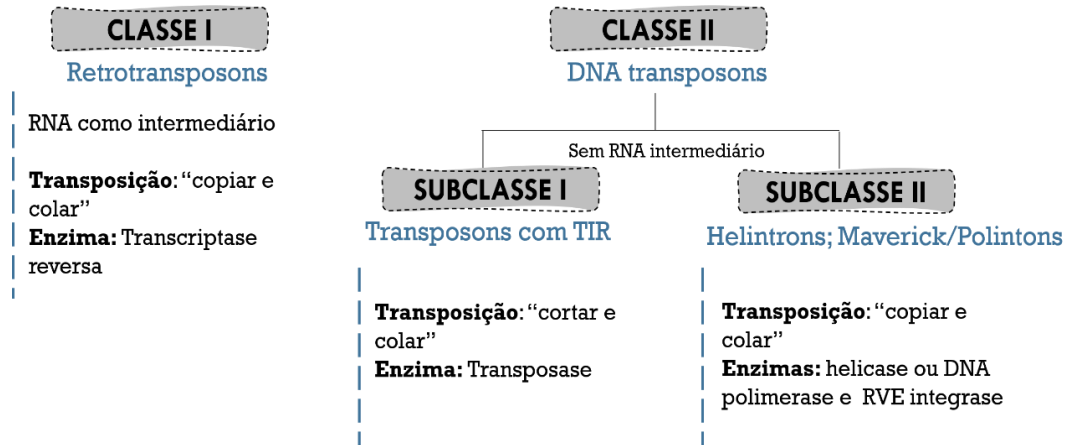


Figura 11 - Esquema ilustrando os níveis mais altos da classificação hierárquica proposta por Wicker e colaboradores (2007) para os transposons. TIR: repetições terminais invertidas (*terminal inverted repeats*)

1.7 Vírus gigantes e seu mobiloma

Usualmente, a relação entre os vírus e os EGMs é caracterizada por considerar os vírus como parte do mobiloma de determinada célula e não por conterem seu próprio mobiloma. (DESNUES *et al.*, 2012). Além disso, devido à ideia clássica de que as partículas e genomas dos vírus são sempre pequenos e compactos, a presença de EGMs em genomas virais é, geralmente, considerada rara (FILÉE, 2018). Entretanto, alguns tipos de elementos genéticos móveis já foram detectados em genomas virais, sendo que boa parte deles são encontrados em vírus pertencentes a famílias de NCLDV (FILÉE, 2018). Como já se sabe, a descoberta dos mimivírus questionou vários paradigmas. No caso dos EGMs isso não foi diferente, pois os mimivírus possuem o seu próprio mobiloma, que chama a atenção por ser bastante complexo e específico (COLSON *et al.*, 2017).

Na família *Mimiviridae*, os primeiros EGMs descritos foram elementos pertencentes ao grupo dos parasitas moleculares *self-splicing*. Foram detectados, no APMV, quatro íntrons *self-splicing* localizados no gene de uma RNA polimerase, bem como uma inteína localizada na proteína DNA polimerase B (RAOULT *et al.*, 2004; OGATA; RAOULT; CLAVERIE, 2005). A complexidade e singularidade do mobiloma dos mimivírus foi evidenciada com a descoberta da capacidade dos virófagos de se integrar no genoma dos mimivírus, formando os provirófagos. Além disso, a descrição dos transpovirons, uma nova classe de elementos móveis, também foi importante, nesse sentido (DESNUES *et al.*, 2012). Além dos mimivírus, a associação de EGMs com o genoma de outros vírus gigantes de ameba também foi descrita nos últimos anos, como é o caso dos pandoravírus. O isolado pandoravírus salinus foi o primeiro

vírus descrito portando um transposon de DNA clássico em seu genoma. Esse transposon foi chamado de *Submariner* e foi proposto que ele pertencesse à superfamília *Tc1/mariner* de DNA transposons da subclasse I (SUN *et al.*, 2015). Dessa forma, considerando o papel evolutivo dos EGMs para seus genomas hospedeiros, o estudo do mobiloma dos vírus gigantes pode contribuir bastante para o melhor entendimento da evolução dos vírus gigantes.

1.7.1 Provírfagos

Há quase uma década, foi descrito um novo isolado de mimivírus, o lentillevírus, e de um vírfago sputnik, chamado de sputnik 2. Após o isolamento desses novos organismos, os pesquisadores tentaram separar as partículas do vírfago das de seu vírus gigante hospedeiro, por meio da inativação do vírfago utilizando calor e dessecação. Curiosamente, eles perceberam que, mesmo após essa inativação, o vírfago continuava liberando suas partículas. Por meio de várias metodologias diferentes, esse trabalho revelou a primeira evidência de que a sequência de um vírfago poderia se integrar no genoma de seu vírus gigante hospedeiro, formando os chamados provírfagos (DESNUES *et al.*, 2012). Assim, foi sugerido que o sputnik 2 utilizasse, como *hot spot* de integração, uma região em comum entre a sua sequência e a do lentillevírus. Essa região em comum contém 395 nucleotídeos e faz parte de genes que codificam proteínas contendo repetição tipo colágeno, tanto no vírfago, quanto no vírus gigante. Apesar disso, também foi proposto que o vírfago poderia se inserir em sítios múltiplos e aleatórios do genoma do lentillevírus (DESNUES *et al.*, 2012).

Embora descritos pela primeira vez no genoma de mimivírus, os provírfagos não são encontrados restritamente nos genomas de mimivírus. Um estudo anterior descreveu elementos similares a vírfagos (VLEs, *virophage-like elements*) integrados ao genoma de algas mixotróficas marinhas da espécie *Bigelowiella natans* (BLANC *et al.*, 2015). Outro trabalho descreveu que os vírfagos mavírus, parasitas do CroV, são capazes de integrar seu genoma em múltiplos sítios do DNA da célula hospedeira, o protista *Cafeteria roenbergensis*. (FISCHER; HACKL, 2016). Os pesquisadores mostraram que os genes do provírfago de mavírus, contidos no genoma do protista, se mantêm inexpressivos, sendo ativados após infecção pelo vírus gigante, o CroV. Essa ativação leva à produção de partículas infecciosas do vírfago mavírus, que, ao infectarem novas células, suprimem a replicação do genoma do CroV, levando ao aumento da sobrevivência da célula hospedeira (FISCHER; HACKL, 2016). Ambos os trabalhos sugerem um papel ecológico dos provírfagos, ao estarem envolvidos em uma relação mutualística com o eucarioto hospedeiro. Eles fornecem para a célula um mecanismo de defesa

antiviral, enquanto recebem em troca a possibilidade de aumentar a sua taxa de coinfeção com um vírus gigante, uma vez que persistem integrados no genoma da célula hospedeira (FISCHER; HACKL, 2016; BLANC *et al.*, 2015).

1.7.2 Transpovirons

Na ocasião da descrição do virófago sputnik 2 e de seu vírus gigante hospedeiro, o lentillevírus, os pesquisadores também descreveram um elemento genético móvel até então desconhecido. Por apresentar características similares a de um transposon, o novo EGM foi denominado transpoviron, uma vez que foi detectado no genoma de um vírus (DESNUES *et al.*, 2012). Os transpovirons são EGMs que consistem em uma sequência de DNA linear extracromossômica, formada por cerca de 7 kpb, capazes de codificar entre 6 e 8 ORFs e que, geralmente, apresentam repetições terminais invertidas em suas extremidades. Eles podem se inserir aleatoriamente no genoma do vírus gigante e, com menos frequência, no genoma de virófagos (DESNUES *et al.* 2012). Além disso, a sequência de DNA do transpoviron, bem como proteínas específicas codificadas por ela, também podem ser encontradas dentro das partículas de vírus gigante e de virófagos (DESNUES *et al.*, 2012; JEUDY *et al.*, 2019).

Além do transpoviron associado ao lentillevírus, que foi o primeiro a ser descoberto, as análises de sequências de outros mimivírus já depositadas em bancos de dados permitiram, inicialmente, a identificação de outros 3 transpovirons, cada um associado a um vírus pertencente às linhagens A, B e C de mimivírus, sendo eles o mamavírus, o moumouvírus monve e o megavírus courdo 7, respectivamente (DESNUES *et al.*, 2012). Recentemente, outros dois transpovirons foram descritos, um em associação a um mimivírus da linhagem B (moumouvírus australiensis) e outro em associação a um mimivírus da linhagem C (megavírus vitis) (JEUDY *et al.*, 2019). Uma lista sumarizando os seis transpovirons já descritos pode ser observada na tabela 3.

Tabela 3 – Resumo dos seis transpovirons já descritos e os seus mimivírus associados.

Transpoviron (abreviação)	Mimivírus associado (linhagem)	Referência
Transpoviron do lentillevírus (Tvlent)	Lentillevírus (A)	DESNUES <i>et al.</i> , 2012
Transpoviron do mamavírus (Tvmama)	Mamavírus (A)	DESNUES <i>et al.</i> , 2012
Transpoviron do moumouvírus monve (Tvmonve)	Moumouvírus monve (B)	DESNUES <i>et al.</i> , 2012
Transpoviron do moumouvírus australiensis (Tvmaust)	Moumouvírus australiensis (B)	JEUDY <i>et al.</i> , 2019
Transpoviron do megavírus courdo7 (Tvcourdo)	Megavírus courdo7 (C)	DESNUES <i>et al.</i> , 2012
Transpoviron do megavírus vitis (Tvmvitis)	Megavírus vitis (C)	JEUDY <i>et al.</i> , 2019

No que diz respeito ao mecanismo de replicação dos transpovirons, embora ainda sejam necessários estudos aprofundados para elucidar esse processo, foi descrito que os transpovirons devem utilizar a maquinaria de replicação e a fábrica viral do vírus gigante hospedeiro para se replicar, uma vez que não codificam enzimas próprias, como a DNA polimerase (DESNUES *et al.*, 2012; JEUDY *et al.*, 2019). A análise comparativa do genoma dos transpovirons revela que três genes codificam proteínas que estão presentes em todos os transpovirons, representando seus *core genes*. Uma delas é uma proteína com função desconhecida, enquanto as outras duas são uma proteína com possível função de helicase e uma proteína com domínio de dedo de zinco (JEUDY *et al.*, 2019). Além disso, foi descrito que todos os transpovirons compartilham uma pequena proteína contendo um segmento transmembrana e que pelo menos 7 proteínas são compartilhadas por pelo menos dois transpovirons. Apesar disso, muitos dos genes preditos para os transpovirons codificam proteínas com função ainda desconhecida. Curiosamente, as análises filogenéticas das proteínas codificadas pelos três *core genes* de transpovirons, mostraram que eles se agrupam em três clados distintos, os quais refletem a divisão das três linhagens dos mimivírus, de acordo com o vírus associado a cada transpoviron (Figura 12).

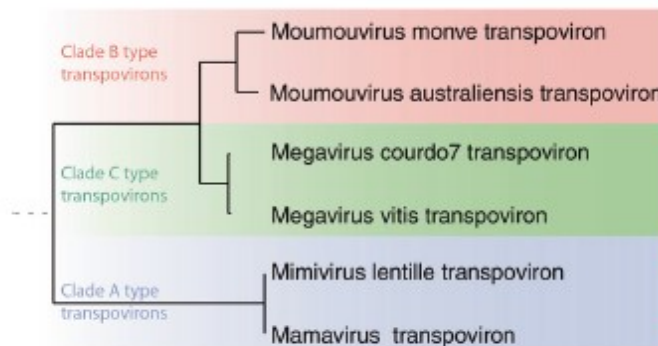


Figura 12 - Árvore filogenética de sequências concatenadas das proteínas codificadas pelos *core genes* dos transpovirons. Fonte: Adaptado de Jeudy *et al.* 2019

Alguns dos genes codificados pelos transpovirons, como a proteína contendo dedo de zinco, apresentam homologia com genes de virófagos, o que indica que os genes de transpoviron e de virófago podem se recombinar, já que compõem o mobiloma do vírus gigante (DESNUES *et al.*, 2012). Essa relação entre transpovirons, virófagos e vírus gigantes parece ser bastante próxima e outra característica que reforça essa ideia é o fato de os virófagos também funcionarem como veículos de propagação dos transpovirons entre os diferentes vírus gigantes. (JEUDY *et al.*, 2019)

1.8 O papel do mobiloma na evolução dos vírus gigantes e virófagos

O genoma dos vírus gigantes é caracterizado por possuir uma considerável quantidade de genes, possivelmente, adquiridos horizontalmente a partir de diferentes organismos, como bactérias, fungos, arqueias e até mesmo outros vírus (COLSON *et al.*, 2017). O genoma dos virófagos também é caracterizado pelo seu mosaicismos, contendo genes ligados evolutivamente a vírus que infectam organismos pertencentes aos três domínios da vida (LA SCOLA *et al.*, 2008). Propõe-se que esse tipo de mosaicismos genético esteja relacionado ao fato de as amebas hospedeiras serem capazes de fagocitar diferentes organismos, como bactérias e outros vírus. Assim, elas funcionam como *hot spots* para essa intensa transferência horizontal de genes e favorecem o surgimento de diferentes repertórios genéticos. (BOYER *et al.*, 2009; RAOULT; BOYER, 2010).

Essa intensa transferência de genes também pode ser feita com os elementos genéticos móveis. A descoberta dos virófagos foi essencial na formação de conexões evolutivas entre os vírus gigantes e os EGMs. A descrição do virófago mavírus e a análise de seus genes revelaram que esses virófagos possuem uma relação evolutiva com DNA transposons (classe II), do tipo *Maverick/Polinton*. O mavírus e os polintons apresentam tamanho e estruturas de genoma similares, além de compartilharem algumas proteínas homólogas, como uma integrase e uma DNA polimerase (FISCHER; SUTTLE, 2011). Os polintons estão distribuídos pelos genomas de diversos eucariotos, o que sugere que eles tenham tido sucesso evolutivo e uma origem bem antiga (KOONIN; KRUPOVIC; YUTIN, 2015).

Além disso, estudos anteriores mostraram que alguns polintons codificam duas proteínas de capsídeo que são essenciais para a morfogênese de partículas virais, uma MCP com estrutura DJR e uma mCP (KRUPOVIC; BAMFORD; KOONIN, 2014). Assim, os polintons que codificam proteínas de capsídeo são chamados de polintovírus. Esses achados foram reforçados pela detecção de um grupo de polintovírus em análises de metagenômica, utilizando a MCP de um virófago como base (YUTIN *et al.*, 2015). Junto às proteínas de capsídeo detectadas, os polintons também podem codificar uma ATPase relacionada ao empacotamento de genoma e uma protease relacionada com a maturação do virion, formando um conjunto de proteínas envolvidas na morfogênese. Apesar de ainda não terem sido identificados em cultura, isso sugere que esses EGMs possam produzir vírions e infectar novos hospedeiros sob determinadas condições, tendo um estilo de vida duplo, que combina características de vírus e de transposons. (KOONIN; KRUPOVIC; YUTIN, 2015)

As análises filogenéticas de genes conservados em *Nucleocytoviricota* evidenciam a relação entre os vírus gigantes e vírus menores. Adicionalmente, estudos de genômica comparativa apontam que os polintons e polintovírus são protagonistas de um possível cenário evolutivo, no qual eles teriam se originado de pequenos bacteriófagos da família *Tectiviridae*, durante o surgimento dos eucariotos, e, posteriormente, teriam dado origem aos virófagos e aos NCLDV (KOONIN; KRUPOVIC; YUTIN, 2015). Essas relações evolutivas são reflexo de uma rede de interações, com conjuntos de genes compartilhados entre diferentes grupos de vírus, virófagos e elementos genéticos móveis (YUTIN; RAOULT; KOONIN, 2013). A presença das principais classes de elementos genéticos móveis no genoma dos vírus gigantes pode ser, assim, considerada uma característica essencial para a evolução, já que esse mobiloma pode funcionar como uma plataforma de TGH e, conseqüentemente, de adaptação e de aquisição de novas características (DESNUES *et al.*, 2012).

2 JUSTIFICATIVA

Notavelmente, os vírus gigantes atraíram a atenção da comunidade científica devido a sua singularidade. Desde a descoberta dos mimivírus, diversas pesquisas bem-sucedidas vêm sendo feitas, com o objetivo de obter novos isolados e caracterizar os seus mais diferentes aspectos, o que tem levado a várias descobertas instigantes. Dentre essas descobertas está o fato de os vírus gigante serem infectados por vírus menores chamados de virófagos, além de possuírem o seu próprio mobiloma, composto por diferentes elementos genéticos móveis (EGMs). O mobiloma dos mimivírus é composto por íntrons e inteínas *self-splicing*, além de outros dois elementos diferenciados chamados de provirófagos e transpovirons.

A relação entre os vírus gigantes, virófagos e transpovirons parece ser bem próxima, uma vez que, tanto os virófagos, quanto os transpovirons são dependentes da maquinaria do vírus gigante hospedeiro para expressarem os seus genes e se replicarem. A combinação entre três diferentes genomas, sendo replicados em uma mesma fábrica viral, somada à presença de diferentes microrganismos fagocitados pela ameba hospedeira, torna esse ambiente celular um importante *hot spot* para a transferência gênica horizontal (TGH). Assim, a dependência da maquinaria viral para a replicação dos transpovirons e virófagos, bem com a presença de genes homólogos entre eles, evidencia a importância desses elementos para a TGH, tornando-os grandes impulsionadores da evolução dos vírus gigantes. Além disso, a possibilidade de integração do genoma do virófago no DNA do vírus gigante, bem como a associação dos transpovirons a diferentes linhagens de mimivírus, também representam interações instigantes, para as quais alguns mecanismos são ainda pouco conhecidos, já que os provirófagos e transpovirons ainda são pouco estudados.

Após a coleta de amostras de uma lagoa urbana brasileira, um novo isolado de vírus gigante foi obtido pelo nosso grupo. Análises preliminares propuseram que esse novo isolado fosse portador de um mobiloma próprio e típico de mimivírus, composto por um provirófago e um transpoviron. Pensando nisso, a caracterização genômica desse novo isolado, junto a de seus EGMs, pode expandir o conhecimento básico sobre elementos ainda pouco estudados. Esse estudo pode contribuir, assim, para o melhor entendimento da relação entre os vírus gigantes e seu mobiloma, enriquecendo, portanto, o conhecimento acerca da evolução e da origem dos virófagos e dos vírus gigantes.

3 OBJETIVOS

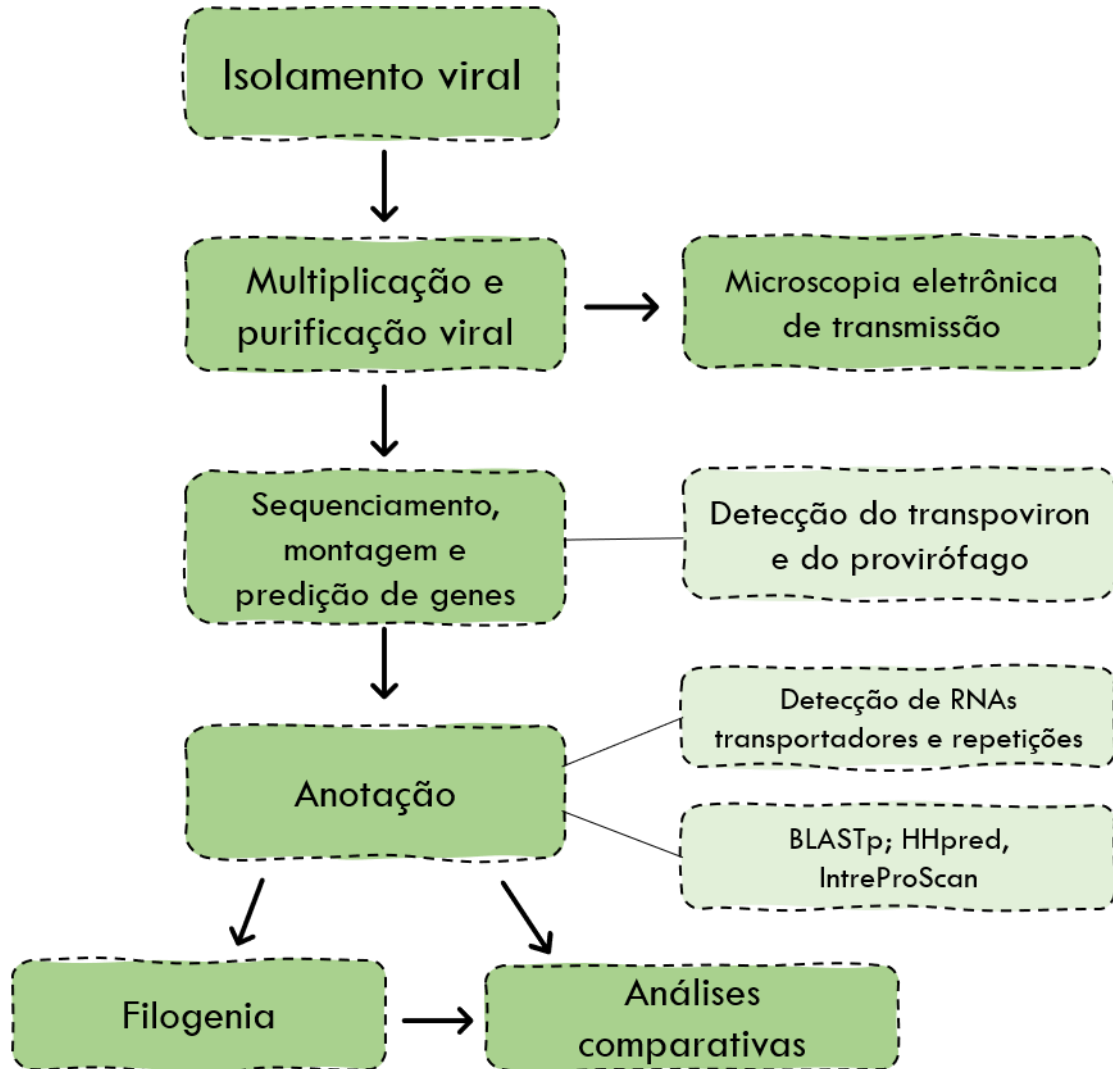
3.1 Objetivo geral

Realizar a identificação e a caracterização genômica e filogenética de um novo isolado de vírus gigante, bem como de seus elementos genéticos móveis associados, um provirófago e um transpoviron.

3.2 Objetivos específicos

- Obter uma visão geral das partículas do vírus gigante, sua fábrica e do virófago associado, bem como da interação de ambas as partículas dentro de células de *Acanthamoeba castellanii*.
- Realizar o sequenciamento, a montagem do genoma do vírus gigante e detectar os elementos genéticos móveis que compõem o mobiloma do vírus gigante isolado, um provirófago e um transpoviron.
- Anotar o genoma do vírus gigante, do virófago e do transpoviron associados, com base em análise de similaridade de sequência de aminoácidos (*BLASTp*) e estrutural (*HHpred*), comparando seus resultados.
- Analisar filogeneticamente as proteínas conservadas preditas para o vírus gigante isolado e seu virófago associado, bem como as diferentes proteínas preditas para o transpoviron.
- Identificar os genes vizinhos da região onde o provirófago está, possivelmente, integrado no genoma do vírus gigante isolado, comparando-os com genes de diferentes mimivírus e virófagos.
- Buscar por possíveis estruturas e ou/funções (*HHpred*) ainda não preditas para as proteínas de todos os transpovirons descritos anteriormente e comparar o genoma do transpoviron detectado com o dos outros transpovirons descritos.

4 FLUXOGRAMA



5 MATERIAIS E MÉTODOS

5.1 Isolamento viral e obtenção de imagens de microscopia dos isolados

5.1.1 Cultivo celular

As células utilizadas durante o trabalho foram amebas da espécie *Acanthamoeba castellanii*, originadas da organização estadunidense *American Type Culture Collection* (ATCC 30234). As células foram cultivadas em garrafas de cultivo celular, contendo meio Peptona-Extrato de levedura-Glicose (PYG - *Peptone-Yeast extract-Glucose*), cujo protocolo de preparo pode ser encontrado no anexo 1 deste trabalho. O meio PYG foi suplementado com um conjunto de antimicrobianos, sendo eles a penicilina (500 U /mL - Schering-Plough, Brasil), a gentamicina (50 mg/mL - Schering-Plough, Brasil) e a anfotericina B (25 mg/mL - Cristalia, São Paulo, Brasil), de modo a evitar possíveis contaminações na cultura.

As células de ameba foram mantidas por meio de subcultivos periódicos, na medida em que eram necessárias para os experimentos. Para tanto, as células eram inoculadas na garrafa contendo o meio PYG, incubadas em estufa a 29°C e o crescimento da monocamada era observado ao longo dos dias, até que a confluência ideal para o subcultivo fosse atingida. Em seguida, a monocamada de células era despreendida mecanicamente e as células contadas em Câmara de Neubauer. O conteúdo era, então, dividido em novas garrafas contendo meio PYG fresco, de acordo com a quantidade de células necessárias para os experimentos. As células subcultivadas foram utilizadas para os experimentos de multiplicação e titulação viral, bem como para o preparo de amostras para visualização em microscopia eletrônica de transmissão.

5.1.2 Isolamento viral

O vírus objeto de estudo deste trabalho foi isolado a partir de amostras de água da Lagoa da Pampulha, em Belo Horizonte, Minas Gerais, que foram coletadas em tubos estéreis, no ano de 2015. Após a coleta, a amostra foi submetida ao método de prospecção. Durante os ensaios de prospecção, 100 microlitros (μL) da amostra coletada foram inoculados em placas de 96 poços, contendo 4×10^4 amebas por poço. Os poços foram observados diariamente, na busca por efeitos citopáticos, e as amostras foram submetidas a três passagens subsequentes (ANDRADE *et al.*, 2018). Os poços que apresentaram efeito citopático tiveram seu conteúdo coletado, o qual foi analisado para confirmação da presença de vírus, utilizando microscopia eletrônica com contrastação negativa.

Nessa análise, foram observadas partículas virais grandes e pequenas, sugestivas de um mimivírus e de um virófago. As partículas virais foram então separadas das partículas de

virófago por meio da inativação deste último utilizando calor (68°C/1h) e dessecação por 48 horas, como previamente descrito (DESNUES et al, 2012). Ensaios posteriores utilizando a amostra contendo apenas a partícula do vírus gigante, já que o virófago havia sido inativado, mostraram que o virófago ainda continuava se multiplicando, sugerindo que seu genoma estivesse supostamente integrado ao do vírus gigante. Uma amostra contendo o mimivírus e o virófago inativado foi preparada no Laboratório de Vírus do ICB/UFMG para, assim, ser utilizada nos estudos deste trabalho. Registros SISGEN: A473BD3 e A702EB8.

5.1.3 Multiplicação e purificação viral

Para o desenvolvimento dos experimentos do trabalho, foi necessário multiplicar e purificar o vírus isolado. A amostra original, com o isolado viral, foi inicialmente utilizada para a produção de um estoque de “pool semente” por meio da inoculação de 30 µl da amostra em garrafas de cultura de célula contendo 8×10^6 amebas *A. castellanii* e meio PYG suplementado com antibióticos. Após a observação do efeito citopático, o sobrenadante foi coletado, alíquotado em microtubos de 1,5 mL e, então, uma porção foi titulada para posterior produção do “pool trabalho”, enquanto a outra foi estocada em freezer a -20°C.

Para a produção do pool trabalho, foram utilizadas garrafas de cultura de vidro (300 cm²), contendo $1,4 \times 10^7$ células de ameba e 35 mL de meio PYG, que foram inoculadas com o pool semente na multiplicidade de infecção (MOI) de 0,01, como sugerido por Abrahão e colegas (2014b). As amostras foram incubadas no equipamento Roller, que manteve as garrafas em lenta rotação, sob a temperatura de 32°C. Após observação do efeito citopático, o sobrenadante de cada garrafa foi coletado em tubos de 50 mL e utilizado, posteriormente, para a purificação viral.

A primeira etapa da execução da purificação viral foi a realização de três sessões de congelamento e descongelamento dos tubos contendo o sobrenadante coletado do “pool trabalho”. O objetivo dessa sessão é promover a lise de células eventualmente não rompidas, permitindo a liberação de partículas virais no sobrenadante. Em seguida, o conteúdo foi delicadamente pipetado sobre um colchão de 10 mL de solução de sacarose a 22% (p/v) que havia sido adicionado no fundo de tubos próprios para ultracentrifuga *Combi Sorvall Rotor AH-629*. Os tubos foram balanceados, utilizando balança de precisão, e submetidos à ultracentrifugação a 14.000 rotações por minuto (rpm) (rotor AH-629) com temperatura entre 4°C e 8°C durante 30 minutos. Posteriormente, o sobrenadante dos tubos ultracentrifugados foram descartados e o *pellet* foi ressuspendido em 500 µL de tampão fosfato salina (PBS) 1x.

Para este trabalho, foram feitas duas rodadas de purificação do “pool trabalho” viral. Uma delas foi alíquotada em criotubo de 1 mL e enviada para o sequenciamento, enquanto a outra foi armazenada em freezer a -20°C, em alíquotas de 10 µL feitas em microtubos de 0,2 mL. Essas alíquotas foram submetidas à titulação e, posteriormente, utilizadas para visualização por microscopia eletrônica de transmissão.

5.1.4 Titulação viral

Os ensaios de titulação viral foram desenvolvidos por meio do método de diluição limitante (*end point*), como descrito por Reed e Muench (1938). Para tanto, foram feitas diluições seriadas em PBS 1x, para uma razão de 10 (10^{-1} a 10^{-11}), a partir da amostra de interesse para a titulação. Posteriormente, um volume de 100 µL de cada diluição era inoculado, em quadriplicata, nos poços de placas de 96 poços previamente inoculadas com 100 µL de meio PYG contendo 4×10^4 células de *Acanthamoeba castellanii* por poço.

Quatro poços das placas eram reservados para o controle celular, ou seja, não eram inoculados com vírus. Finalmente, as placas eram vedadas, incubadas em estufa a 28°C e observadas diariamente, a fim de monitorar o surgimento de efeito citopático. Após um total de quatro dias após a inoculação, era feita a leitura da titulação, com base nos cálculos propostos por Reed e Muench (1938), os quais têm o objetivo de medir a quantidade de vírus necessária para causar efeito citopático em 50% das células inoculadas com vírus. Os resultados foram expressos em TCID₅₀/mL.

5.1.5 Microscopia eletrônica de transmissão

Para caracterizar morfológicamente os isolados, foi realizado o preparo da amostra para visualização em microscopia eletrônica de transmissão (MET). Primeiramente, foi necessário o cultivo de amebas *A. castellanii* em garrafas contendo 7×10^6 células e 25 mL de meio PYG. Em seguida, uma amostra do vírus purificado foi inoculada na garrafa em M.O.I de 0.01 e esta foi observada diariamente para avaliação do efeito citopático. Dessa forma, foi possível observar o ciclo de multiplicação do vírus de forma assincrônica, uma vez que a baixa quantidade de partículas inoculadas por célula faz com que as células sejam infectadas em momentos diferentes após a inoculação.

Após a observação de efeito citopático, o próximo passo foi desprender a monocamada de células mecanicamente e coletar o conteúdo da garrafa em tubos de 50 mL, que foram centrifugados em centrífuga refrigerada *Sorvall RT6000B* a 3.000 rpm por 10 minutos.

Posteriormente à centrifugação, o sobrenadante foi descartado e o *pellet* ressuspensionado em 5 mL de solução de tampão fosfato de sódio a 0,1 M. Este procedimento foi repetido duas vezes. O sobrenadante resultante do processo anterior foi novamente descartado e o *pellet* ressuspensionado, dessa vez com a solução fixadora, produzida a partir de tampão fosfato de sódio 0,1 M e glutaraldeído a 2,5% (p/v). A amostra foi incubada por duas horas, sob temperatura ambiente e rotação, e após esse tempo foi submetida a uma nova centrifugação a 3.000 rpm por 10 minutos, na mesma centrífuga.

Por fim, o sobrenadante foi descartado e o *pellet* ressuspensionado novamente na solução tampão fosfato 0,1M. O conteúdo final foi alíquotado em um microtubo de 1,5mL, que foi centrifugado em centrífuga *eppendorf 5415R* a 3000 rpm por 10 minutos e depois armazenado em geladeira (4°C) até a realização das próximas etapas. O material foi transportado até o Centro de Microscopia da UFMG (CM-UFMG), cujo corpo técnico realizou as preparações posteriores, dentre as quais se incluem a desidratação, fixação secundária, inclusão em resina, contrastação com tetróxido de ósmio e posterior ultramicrotomia. Após o preparo, a amostra foi observada no CM-UFMG, por meio do Microscópio Eletrônico de Transmissão Tecnai G2-12- SpiritBiotwin FEI - 120 kV. Adicionalmente, os tamanhos das partículas observadas na microscopia foram, posteriormente, medidos utilizando o software ImageJ (SCHNEIDER; RASBAND; ELICEIRI, 2012), com base na escala fornecida pelas próprias imagens registradas no CM-UFMG.

5.2 Caracterização genômica dos isolados e do transpoviron

5.2.1 Sequenciamento, montagem e predição de genes

As amostras contendo o vírus purificado foram enviadas para o sequenciamento do genoma na empresa BPI – Biotecnologia, Pesquisa e Inovação, localizada em Botucatu, São Paulo. Foi utilizada a plataforma de sequenciamento *Illumina MiSeq*, com bibliotecas de fragmentos do tipo *paired end*. O controle de qualidade das leituras (*reads*) obtidas após o sequenciamento foi feito utilizando o software *FastQC*.

A montagem *de novo* das *reads* obtidas foi feita por meio do programa *Spades 3.12*, utilizando os seus parâmetros padrão (BANKEVICH *et al.*, 2012; PRJIBELSKI *et al.*, 2020). O *scaffolds* resultantes dessa montagem foram comparados com sequências de banco de dados, utilizando o *BLASTn* (banco de dados: nr/nt; *expect threshold*: 10^{-3}), otimizado para buscar sequências altamente similares (megablast). Em seguida, a sequência indicada como o *best hit* para os *scaffolds* foi utilizada para finalizar a montagem do genoma do vírus gigante com base

em sequência de referência, que foi realizada por meio do software *MeDuSa* online (BOSI *et al.*, 2015). As sequências obtidas foram, posteriormente, submetidas à predição de genes por meio do software *GeneMarkS* (BESEMER *et al.*, 2005). Os genes preditos foram filtrados, mantendo apenas aqueles que codificavam proteínas com mais de 50 aminoácidos. Dessa forma, foi possível obter o número de ORFs, as sequências de nucleotídeos e aminoácidos de cada ORF, além do conteúdo de G-C.

5.2.2 Anotação de genes

A anotação dos genes preditos para os três genomas estudados foi feita por meio de dois métodos principais. Em primeiro lugar, cada uma das sequências de aminoácidos foi analisada utilizando o *BLASTp* (banco de dados: nr; *expect threshold*: 10^{-3}), em busca por sequências similares. Em segundo lugar, as mesmas sequências foram analisadas manualmente no *HHpred*, servidor online que permite a predição de funções e estruturas, com base em homologia. Na análise do *HHpred*, realizada entre outubro/2020 e janeiro/2021, foram utilizados os bancos de dados PDB_mmCIF70, PDB_mmCIF30 e SCOPe70_2.07 e parâmetros padrão (SODING *et al.*, 2005; ZIMMERMANN *et al.*, 2018). Durante as buscas pelo *HHpred*, foram considerados como *hits* válidos aqueles que apresentavam valores de probabilidade maiores do que 95%. Caso a probabilidade fosse menor do que 95%, o *e-value* deveria ser menor ou igual a 1 (SODING *et al.*, 2005). As análises realizadas no *BLASTp* e no *HHpred* foram comparadas entre si para obtenção da anotação final e os gráficos foram elaborados no software *GraphPad Prism for Windows 7.04*. Adicionalmente, as sequências do transpoviron e do viróforo, também tiveram predição de domínios proteicos conservados feita por meio do *InterProScan*, utilizando seus parâmetros padrão (BLUM *et al.*, 2020).

Além disso, para a identificação de regiões codificadoras de RNAt, no genoma do vírus gigante, foram utilizadas as ferramentas ARAGORN e tRNAscanSE (LASLETT *et al.*, 2004; CHAN *et al.*, 2019), a partir de seus parâmetros padrão. A construção de mapas representativos dos genomas estudados foi feita utilizando o servidor *CG view* (STOTHARD; WISHART, 2004) e o software *SnapGene* (da *Insightful Science*; disponível em www.snapgene.com). Já para buscar e analisar regiões genômicas contendo repetições terminais invertidas foram utilizados o programa *Inverted Repeat Finder* (WARBURTON *et al.*, 2004), com parâmetros padrão, e o software *Geneious Prime 2021.0.3* (<https://www.geneious.com>).

5.2.3 Análises genômicas comparativas

Para realizar comparações de valores de identidade de nucleotídeos entre o genoma do virófago isolado e outros virófagos similares, foi realizado um alinhamento, por meio do *BLASTn* (*expect threshold: 10⁻³*), utilizando como sequência de entrada (*query*) o genoma completo do virófago isolado, contra sequências de diferentes virófagos obtidas no *GenBank*. O mesmo método foi utilizado para comparação entre o genoma do transpoviron detectado e outros transpovirons descritos, bem como para comparação entre genes para proteínas contendo repetição de colágeno presentes em diferentes virófagos.

No caso de alinhamentos de proteínas, com o objetivo de obter valores de identidade de aminoácidos, eles foram realizados para comparar proteínas do transpoviron detectado com proteínas de outros transpovirons. As sequências de cada uma das proteínas do transpoviron detectado foram alinhadas, utilizando o *BLASTp* (*expect threshold: 10⁻³*), contra o conjunto de proteínas de cada um dos outros seis transpovirons de maneira a detectar seus homólogos dentro do grupo. Isso também foi feito para proteínas que estavam presentes em outros transpovirons, mas ausentes no transpoviron detectado. O objetivo foi comparar as proteínas do transpoviron detectado com as dos outros transpovirons descritos e realizar uma análise geral sobre a ocorrência dessas proteínas nos transpovirons.

5.3 Análises filogenéticas

As análises filogenéticas foram realizadas para as três diferentes sequências genômicas estudadas neste trabalho. Para filogenia do vírus gigante, foi utilizada a sequência de aminoácidos da proteína DNA polimerase B, já para a sequência de virófago foram utilizadas sequências do conjunto de genes relacionados a morfogênese, que é conservado na sua família (MCP, mCP, ATPase e Protease) e uma sequência de proteína contendo repetição de tripla-hélice de colágeno. As análises filogenéticas realizadas para o transpoviron foram feitas para todas as proteínas que apresentavam sequências homólogas em mais de um clado de transpoviron.

Para a obtenção dos conjuntos de sequências a serem alinhadas, além das proteínas mencionadas anteriormente, foram utilizadas sequências similares a elas, obtidas no banco de dados do NCBI, utilizando o *BLASTp* (banco de dados: nr; *expect threshold: 10⁻³*). Os alinhamentos das sequências foram realizados utilizando o algoritmo *MUSCLE* (EDGAR *et al.*, 2004). Os melhores modelos evolutivos para cada alinhamento foram preditos pelo algoritmo

ModelFinder, a partir do programa *IQtree*, o qual foi utilizado também para a construção das árvores, pelo método de máxima verossimilhança, com 1000 replicatas de *bootstrap* (KALYAANAMOORTHY *et al.*, 2017; NGUYEN *et al.*, 2014). As árvores construídas foram visualizadas e editadas por meio dos programas *MEGA X* e *iTOL* (KUMAR *et al.*, 2018; LETUNIC *et al.*, 2006).

6 RESULTADOS

6.1 Isolamento e imagens de microscopia eletrônica dos isolados

Os ensaios de prospecção permitiram o isolamento de um novo vírus gigante de ameba brasileiro, a partir de amostras de água da Lagoa da Pampulha, em Belo Horizonte - MG. Após o isolamento e análises preliminares, as amostras foram submetidas a análises posteriores, no Laboratório de Vírus da UFMG. A primeira dessas análises foi feita utilizando a microscopia eletrônica de transmissão, cujo objetivo foi identificar o vírus a partir de suas características morfológicas, bem como ter uma visão geral do seu ciclo de multiplicação.

A análise das imagens contendo células de ameba infectadas de maneira assincrônica (M.O.I 0,01) indicou a presença de partículas e fábricas virais com morfologia típica de mimivírus. Quando comparadas, as partículas de mimivírus observadas são compostas por capsídeos com tamanho médio de 419 nm, cercados por uma camada de fibrilas cuja média de tamanho é de 120 nm. Juntos, o capsídeo e as fibrilas compõem partículas com média de 673 nm (Figura 13 A). Além das partículas de mimivírus, nós também observamos pequenas partículas icosaédricas, com tamanho médio de 47 nm, sugestivas de virófago (Figura 13 B). A estimativa das dimensões das partículas foi feita com base na média do tamanho de 50 e 60 partículas de mimivírus e de virófagos, respectivamente.

As partículas de virófago foram observadas isoladas, agrupadas, dentro de vesículas ou associadas a partículas maduras de mimivírus (Figura 13 C-D, setas pretas). Também foi possível observar partículas de virófagos em formação, associadas à fábrica viral do mimivírus, o que sugere a dependência pela maquinaria do vírus gigante, uma característica marcante da família *Lavidaviridae* (Figura 13 E, setas pretas). Além disso, foram observadas algumas partículas defeituosas de mimivírus, fato que pode estar associado à presença de virófagos em suas fábricas virais durante o ciclo (Figura 13 D-E, setas vermelhas).

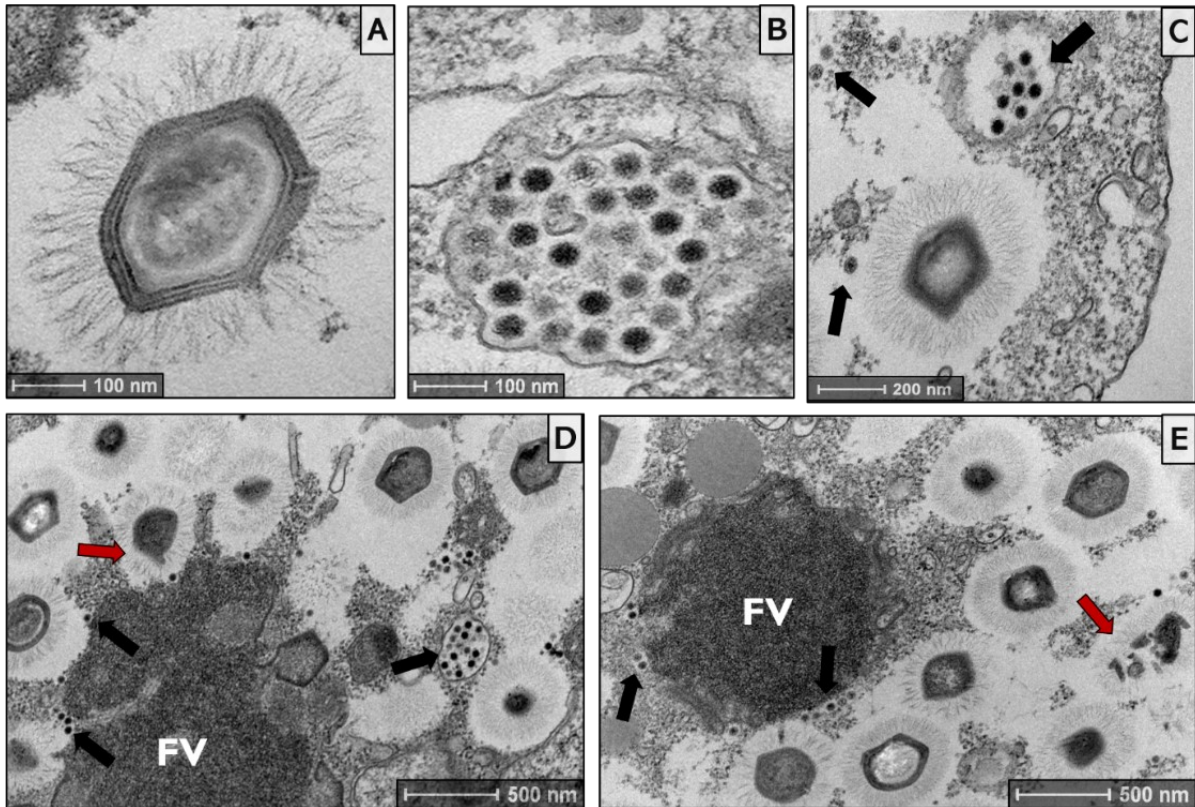


Figura 13 - Imagens de microscopia eletrônica de transmissão mostrando uma visão geral dos novos isolados de mimivírus e de virófago dentro da ameba. (A) Partícula (~673 nm) típica de mimivírus, contendo capsídeo cercado por uma camada de fibrilas. **(B)** Vesícula contendo partículas icosaédricas (~47 nm) típicas de virófago. **(C)** Partículas de virófagos (setas pretas) próximas a mimivírus, isoladas ou agrupadas. **(D-E)** Fábrica viral (FV) típica de mimivírus, onde se pode observar virófagos associados (setas pretas), além de partículas de mimivírus em formação. As setas vermelhas indicam partículas defeituosas de mimivírus.

6.2 Sequenciamento e montagem dos genomas

O sequenciamento das amostras contendo o vírus purificado foi capaz de gerar 281.924 *reads*, com tamanhos que variavam entre 35 e 151 nucleotídeos. Após a análise de controle de qualidade, as *reads* foram submetidas à montagem *de novo*, a qual resultou em 550 *scaffolds*. Essas sequências, geradas após a montagem, foram comparadas com bancos de dados utilizando o *BLASTn*. Essa análise mostrou que grande parte das sequências geradas apresentava alta similaridade com o *Acanthamoeba castellanii* mamavírus. Entretanto, dentre esses *scaffolds*, dois apresentaram uma alta profundidade de sequenciamento e uma similaridade diferente. O primeiro deles, com tamanho de 18.415 pb (profundidade do sequenciamento: 198x), apresentou similaridade com a sequência do virófago sputnik 2. O outro *scaffold* tinha tamanho de 6.823 pb (profundidade do sequenciamento: 78x) e a busca no *BLASTn* mostrou que essa sequência é muito similar à do transpoviron do lentille vírus. Após curadoria manual dos *gaps* presentes nessa última sequência, ela passou a apresentar 6.624 pb.

Os *scaffolds* obtidos após a montagem *de novo* foram, então, organizados por meio da plataforma MeDuSa online, utilizando como referência a sequência de DNA do mamavírus, obtida no GenBank/NCBI (número de acesso: JF801956.1). Assim, as sequências similares ao mamavírus foram montadas em um *scaffold* final de 1.203.963 pb que representa a sequência completa de um isolado de mimivírus brasileiro, que será chamado de mimivírus argentum ao longo deste trabalho. A sequência de virófago, por ser geneticamente similar aos virófangos sputnik, será referida como sputnik argentum, sendo este o terceiro virófago isolado a partir de amostras brasileiras. Já a sequência de transpoviron detectada, será aqui referida como transpoviron do mimivirus argentum (Tvmarg). Dessa forma, obtivemos os três diferentes genomas que foram utilizados como objeto de estudo neste trabalho (Tabela 4)

Tabela 4 - Caracterização no *BLASTn* das três diferentes sequências estudadas neste trabalho, obtidas após o sequenciamento e montagem.

Sequência	Tamanho (pb)	Best hit (<i>BLASTn</i>)	Query cover	e-value	Identidade
Mimivirus argentum	1.203.963	Acanthamoeba castellanii mamavirus strain Hal-V	97%	0.0	99.90%
Sputnik argentum	18.415	Sputnik virophage 2	100.00%	0.0	100.00%
Transpoviron do mimivírus argentum (Tvmarg)	6.624	Mimivirus lentille transpoviron Lentille	100.00%	0.0	99.21%

6.3 Caracterização genômica do mimivirus argentum

6.3.1 Anotação dos genes

A sequência que compõe o genoma do mimivírus argentum é formada por 1.203.963 pb, cujo conteúdo de G-C é de 28,05%. A predição de genes revelou a presença de 1.164 ORFs, com tamanhos que variam entre 156 e 5.799 nucleotídeos. As ORFs estão distribuídas em ambas as fitas de DNA, sendo 574 na fita senso e 590 na fita antisenso. Esses genes codificam proteínas que contém de 51 a 1.932 aminoácidos, as quais foram submetidas à busca por sequências similares, bem como à predição de função e estruturas, por meio do *BLASTp* e do *HHpred*.

A análise inicial das sequências de aminoácidos, realizada no *BLASTp*, revelou que a maioria das proteínas do mimivírus argentum é similar às do mamavírus e do mimivirus (APMV). Dentre os *best hits* do *BLASTp* obtidos para as 1.164 proteínas preditas, 560 foram proteínas do mamavírus (48%), enquanto 444 foram proteínas do APMV (38%). Isso representa

86% do conteúdo de proteínas do mimivírus *argentum* (Figura 14). Para 11% das proteínas, os *best hits* foram sequências de outros mimivírus da linhagem A, como o lentillevírus, hirudovírus, sambavírus e niemeyervírus, além de alguns mimivírus da linhagem C (Figura 14).

Além disso, um fato que chamou nossa atenção durante essas análises, foi a presença de 20 ORFs seguidas (proteínas M888 a M907), cujos *best hits* do *BLASTp* eram proteínas do virófago sputnik, sugerindo a presença de um provirófago completo no genoma do mimivírus *argentum*. Todos os *best hits* do *BLASTp* para cada uma das proteínas do mimivírus *argentum* podem ser encontrados no anexo 2.

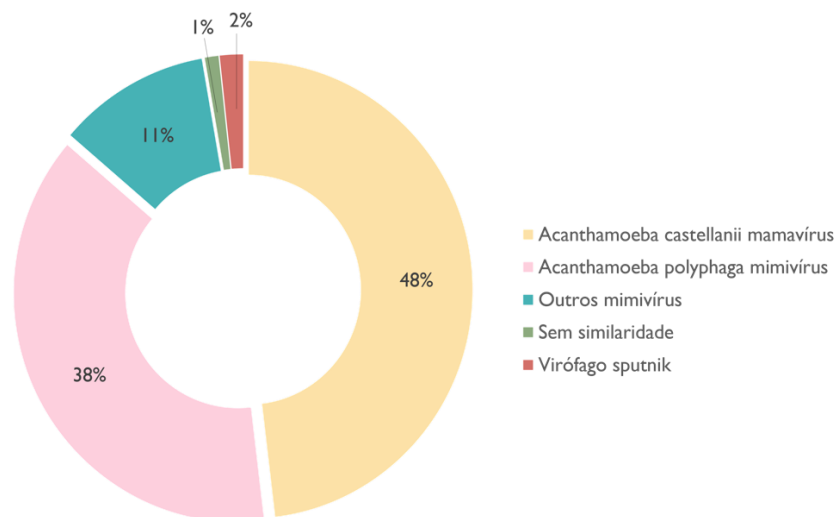


Figura 14 - Caracterização dos *best hits* obtidos no *BLASTp*, durante análise das proteínas do mimivírus *argentum*, em busca por similaridade em bancos de dados.

Após as análises no *BLASTp*, foi possível, então, encontrar sequências similares para quase todas as proteínas do mimivírus *argentum*, com exceção de 12 (1%) que não apresentaram similaridade com o banco de dados de sequências de proteínas não-redundantes (nr) do NCBI (Figura 14). Isso indica que os genes que codificam essas 12 proteínas podem ser considerados ORFans, exclusivas do mimivírus *argentum*. Entretanto, o número total de proteínas para as quais os *best hits* foram “proteínas hipotéticas” foi bastante expressivo. Essas proteínas anotadas como hipotéticas são aquelas para as quais não foram detectadas sequências homólogas ou similares em nenhum banco de dados, na ocasião de sua primeira descrição, compondo o extenso conjunto de ORFans dos mimivírus. Somadas às 12 ORFans exclusivas do mimivírus *argentum*, a quantidade de “proteínas hipotéticas” detectada representou 55,8%

(649/1.164) de todo o conteúdo proteico do novo vírus isolado. Por causa disso, todas as proteínas do mimivírus *argentum* foram analisadas no *HHpred*, com o objetivo de aprimorar a anotação e prever funções e/ou estrutura a partir de outra metodologia e bancos de dados, visando diminuir, assim, a quantidade de “proteínas hipotéticas” que seriam anotadas, caso fossem utilizados apenas os *best hits* do *BLASTp* como base. Os *best hits* obtidos no *HHpred* para cada proteína do mimivírus *argentum* estão disponíveis no anexo 2.

A busca realizada no *HHpred* permitiu a detecção de sequências similares para 197 das 649 proteínas do mimivírus *argentum* que seriam anotadas como “proteínas hipotéticas”, se considerados os *best hits* do *BLASTp*. Por outro lado, as 12 proteínas para as quais não se havia detectado nenhuma sequência similar no *BLASTp* não apresentaram sequências similares nas buscas realizadas no *HHpred*. Dentre as 197 proteínas hipotéticas para as quais foram detectados *hits* no *HHpred*, a maioria apresentou valores de probabilidade considerados altos (>95%). Ao analisar os 197 *best hits* obtidos no *HHpred* para “proteínas hipotéticas”, foi possível categorizar as proteínas correspondentes a cada um deles em seis diferentes grupos: proteínas contendo repetições ou domínios diversos, enzimas diversas, proteínas relacionadas com a transcrição, proteínas estruturais, proteínas relacionadas à replicação e reparo de DNA e proteínas relacionadas à tradução (Figura 15 A).

A maior parte das proteínas encontradas nessa análise (90/197) foi incluída no grupo de proteínas que contém repetições ou domínios diversos. Esse grupo de proteínas é caracterizado por apresentar algumas repetições ou domínios estruturais específicos, cujas funções ainda não são bem determinadas ou são mais generalistas, como a participação em interações proteína-proteína e a ligação a ácidos nucleicos. Ao especificar essas repetições ou domínios detectados, é possível notar que a repetição de anquirina representou a maioria dos *hits*, com 41% (37/90), seguida por 17 *hits* que eram proteínas contendo repetições ricas em leucina (LRR) e/ou repetição FNIP (fenilalanina-asparagina-isoleucina-prolina) e F-box (Figura 15 B). Em terceiro lugar, estão 13 *hits* com proteínas contendo domínios que atuam na ligação com DNA ou RNA, representados, principalmente por proteínas contendo dedo de zinco. Uma outra parte significativa dos *hits* obtidos eram proteínas contendo a repetição BTB/POZ e/ou a repetição kelch, com 6 *hits*, além de proteínas contendo a repetição WD40 e a repetição MORN, com 3 *hits*, e proteína contendo a repetição NHL, com 2 *hits* (Figura 15 B).

Seguindo as proteínas com repetições ou domínios específicos está a categoria de enzimas diversas, que apresentou 55 *hits* (Figura 15 A). Dentre as proteínas obtidas nesses 55 *hits* é

possível encontrar enzimas que desempenham funções diversas. A maior parte delas são enzimas com atividade de transferase, com 13 *hits*, seguida por enzimas com atividade de quinase ou fosfatase, com 10 *hits*, e por enzimas com atividade de nuclease e helicase, com 7 *hits* cada (Figura 15 C). Outros tipos de enzimas também foram detectados em menor quantidade, como aquelas com função de oxidorreductase, protease ou peptidase, ligase, hidrolase e ATPase (Figura 15 C).

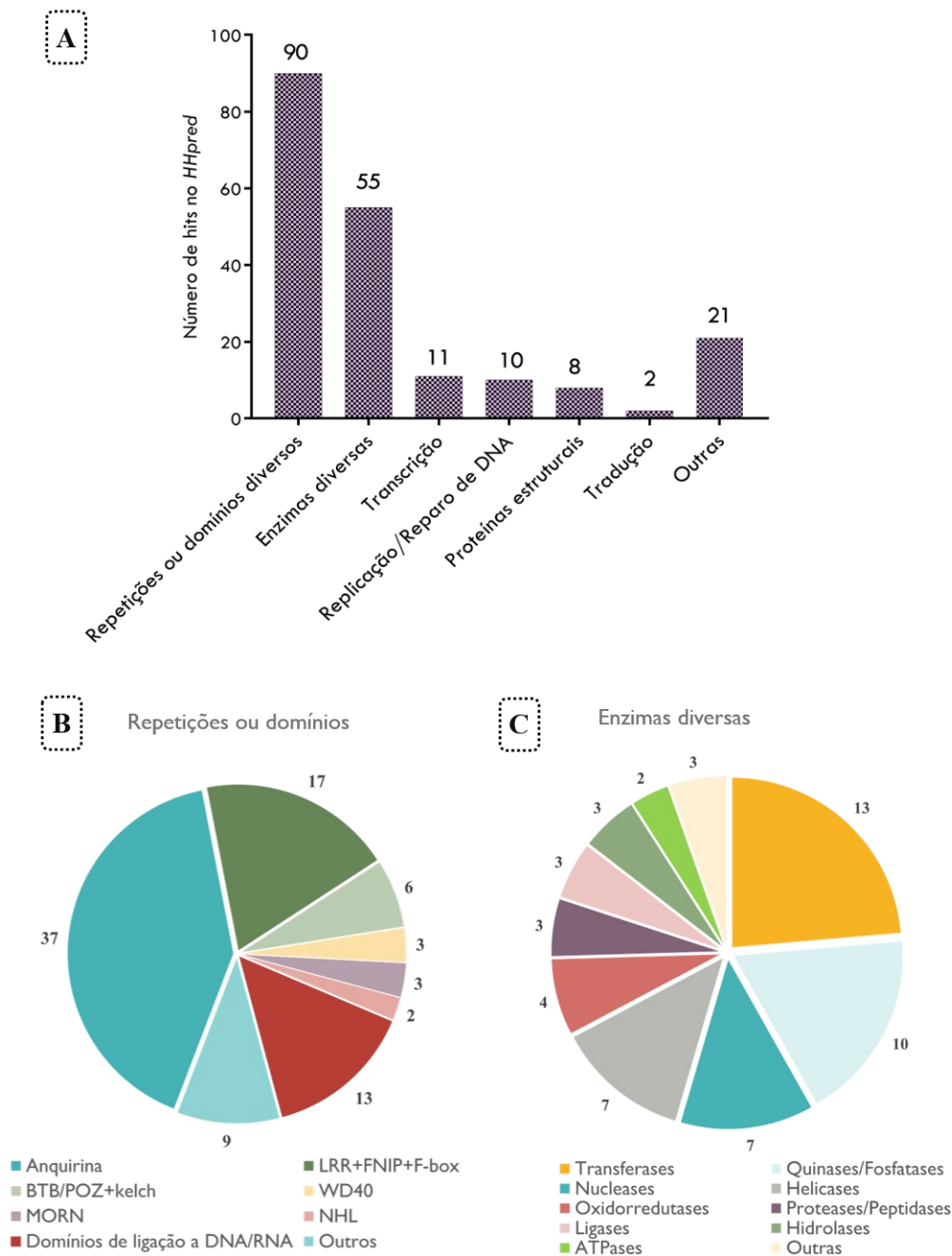


Figura 15 - Resultados das análises no HHpred para as proteínas do mimivírus argentum, cujos best hits do BLASTp eram proteínas hipotéticas (ORFans). (A) Gráfico que apresenta a quantidade de hits obtidos

para cada categoria de proteína indicada, após análise das “proteínas hipotéticas” do mimivírus *argentum* no *HHpred*. (n=197) (B) Gráfico que especifica os 90 *hits* obtidos para diferentes tipos de repetições ou domínios específicos detectados em “proteínas hipotéticas” do mimivírus *argentum*. (C) Detalhamento dos 55 *hits* obtidos no *HHpred* para “proteínas hipotéticas” do mimivírus *argentum* e que correspondem a diferentes tipos de enzimas.

Além das duas categorias já citadas, o restante dos *hits* obtidos para as proteínas hipotéticas do mimivírus *argentum* correspondiam a proteínas que podem estar envolvidas em processos importantes, com a transcrição, a replicação e o reparo de DNA, a tradução e a composição da estrutura do capsídeo viral, como pode ser observado na figura 15 A. Dentre os 11 *hits* que eram proteínas, possivelmente, envolvidas na transcrição e no metabolismo de RNA, estão subunidades de RNA polimerases dependentes de DNA, subunidades de polimerase poli-A e uma enzima envolvida no processo de decapagem do RNA mensageiro. Já os 10 *hits* correspondentes a proteínas envolvidas na replicação ou reparo do DNA incluem subunidades de DNA primase e fatores de replicação, bem como proteínas tipo MuT e uma DNA glicosilase. Com relação aos 8 *hits* em proteínas estruturais, eles correspondem a proteínas de capsídeo, enquanto os 2 *hits* em proteínas, possivelmente, envolvidas na tradução são fatores de iniciação e de liberação de cadeia peptídica.

Dessa forma, para se obter a anotação final dos genes do mimivírus *argentum*, foi necessário comparar os resultados obtidos no *BLASTp*, com os resultados obtidos no *HHpred*. Nessa comparação, o conteúdo total de “proteínas hipotéticas” caiu de 55,8% (649/1.164) para 39% (452/1.164) do total de proteínas do mimivírus *argentum*, mesmo ao acrescentar as 12 novas proteínas hipotéticas exclusivas do novo isolado. Isso representa uma redução aproximada de 17 pontos percentuais no número de proteínas hipotéticas do mimivírus *argentum* (Figura 16). Por outro lado, a obtenção dos 197 *hits* do *HHpred* para proteínas hipotéticas levou, conseqüentemente, a um aumento no número de proteínas de cada uma das categorias propostas. Por exemplo, ao analisar as proteínas do mimivírus *argentum* por meio do *BLASTp*, a quantidade de proteínas contendo as repetições ou domínios específicos mencionados anteriormente, correspondia a cerca de 16% (190/1.164) do total de proteínas do vírus. Com a predição de estruturas e/ou funções no *HHpred*, essa proporção passou a ser de 24% (281/1.164) (Figura 16), dos quais 50,7% (142/281) são proteínas contendo repetição de anquirina. Esse aumento na quantidade de proteínas também pode ser observado para todas as outras categorias, com exceção das 20 proteínas altamente similares a virófago, que nesse caso, foram consideradas uma categoria à parte, independentemente da anotação obtida para cada uma de suas 20 ORFs.

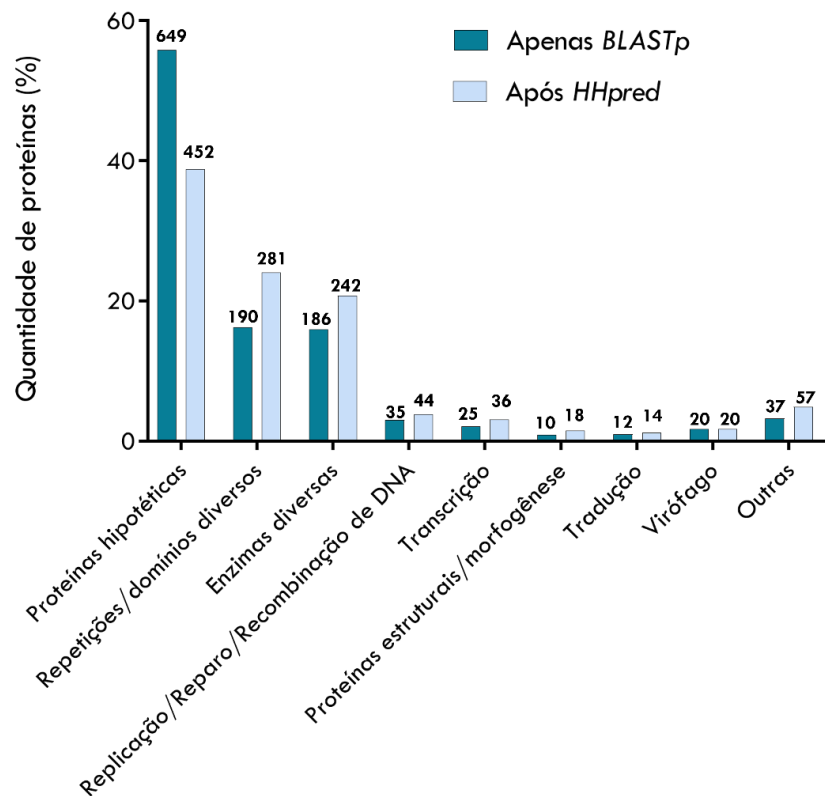


Figura 16 - Comparação entre os dados obtidos após a análise do conjunto de proteínas do mimivírus *argentum* no *BLASTp* e no *HHpred*. O gráfico apresenta, no eixo Y, a porcentagem de proteínas do mimivírus *argentum* que compõe cada uma das categorias indicadas no eixo X, considerando apenas a análise no *BLASTp* ou considerando a análise no *HHpred*. Na extremidade superior de cada uma das barras é possível observar o número de proteínas correspondente a elas. n=1.164

Em contraste com as diferenças mencionadas acima, quando se compara a anotação baseada apenas no *BLASTp* com a anotação baseada também no *HHpred*, é possível observar que a busca no *HHpred* não modificou a anotação da maior parte das proteínas do mimivírus *argentum* e que a maioria das mudanças observadas dizem respeito apenas a diferentes níveis de classificação das proteínas. À exceção das 197 “proteínas hipotéticas” que apresentaram *hits* no *HHpred* e tiveram a sua anotação modificada, apenas 13 proteínas do mimivírus *argentum* apresentaram resultados diferentes ao comparar as duas análises. Dentre essas 13 proteínas, oito apresentaram diferenças no tipo de repetição/domínio detectado, pois, enquanto no *BLASTp* se obteve *hits* para proteínas contendo o domínio BTB/POZ ou repetição de anquirina, no *HHpred*

foram encontrados *hits* para proteínas contendo a repetição WD40 ou LRR/domínio F-box. As diferenças de anotação observadas para as 13 proteínas após as análises no *BLASTp* e no *HHpred* podem ser observadas na Tabela 5.

Tabela 5 – Diferenças entre *best hits* obtidos no *BLASTp* e no *HHpred* para algumas proteínas do mimivírus *argentum*.

Proteína do mimivírus <i>argentum</i>	<i>Best hit</i> no <i>BLASTp</i>	<i>Best hit</i> no <i>HHpred</i>
M90	C-terminal divergent BTB domain-containing protein	WD40 repeat-containing protein
M101	putative BTB/POZ domain-containing protein	WD40 repeat-containing protein
M150	putative ankyrin repeat protein	F-box domain e LRR containing protein
M162	putative BTB/POZ domain-containing protein	WD40 repeat-containing protein
M172	putative BTB/POZ domain-containing protein	WD40 repeat-containing protein
M183	putative F-box protein	Poly A polymerase head domain-containing protein
M367	Proline rich protein	CI-2 family of serine protease inhibitors (serpin)
M393	RING-finger-containing E3 ubiquitin ligase	Helicase conserved C-terminal and RING finger domain
M436	putative DNA repair protein	SMC domain-containing protein
M444	transcription initiation factor TFIIIB	Putative Nucleotidyltransferase
M1019	BTB domain-containing protein	WD40 repeat-containing protein
M1024	putative BTB/POZ domain-containing protein	WD40 repeat-containing protein
M1039	putative BTB/POZ domain-containing protein	WD40 repeat-containing protein

Enfim, a anotação final do genoma do mimivírus *argentum* permitiu a identificação de genes conservados dentre membros dos NCLDV, como é o caso daqueles que codificam a DNA polimerase B (M731), a ATPase de empacotamento A32 (M599) e a DNA primase/helicase D5 (M871). Também foi possível identificar, no mimivírus *argentum*, vários genes típicos de mimivírus da linhagem A, que codificam proteínas envolvidas em diferentes funções durante o ciclo de multiplicação do vírus. O mimivírus *argentum* codifica algumas proteínas envolvidas no reparo do DNA já descritas para mimivírus, como aquelas da família MutS, endonucleases específicas do tipo VII, IV e uvdE, DNA glicosilases, dentre outras. Também foram identificadas no mimivírus *argentum* proteínas do tipo DnaJ, HSP70 e peptidilprolil isomerase (ciclofilina) preditas para atuar como chaperonas. O vírus também apresenta genes para enzimas como a acetilcolinesterase e a glutamina sintase, que foram propostas para estarem envolvidas no metabolismo de lipídeos e de aminoácidos, respectivamente. Além disso, também foi possível observar a presença de genes para proteínas que podem estar envolvidas na glicosilação de proteínas, como vários tipos de glicosiltransferases e a enzima procolágeno, 2-oxoglutarato dioxygenase, para a qual foi proposta a participação na glicosilação de proteínas do tipo colágeno.

Assim como outros mimivírus, o mimivírus *argentum* também codifica um conjunto de genes envolvidos na transcrição e na tradução. Os genes relacionados à transcrição codificam além de subunidade de proteínas como RNA polimerases dependentes de DNA, vários fatores de transcrição. Já os genes envolvidos na tradução de proteínas codificam fatores de tradução de cinco tipos, incluindo os fatores de iniciação IF4A, IF4E e SUI1, fator de alongação eF-Tu e o fator de liberação de cadeia eRF1. Além disso, quatro tipos de aminoacil t-RNA sintetases também são codificados, sendo eles arginil, metionil, cisteinil e tirosil t-RNA sintetases. Ainda como parte do aparato de tradução do mimivírus *argentum*, foram identificados 6 RNAs transportadores, indicados na figura 17. Dessa forma, é possível concluir que, de maneira geral o mimivírus *argentum* apresenta um conteúdo de genes e proteínas muito similar ao dos outros mimivírus, especialmente aqueles pertencentes à linhagem A.

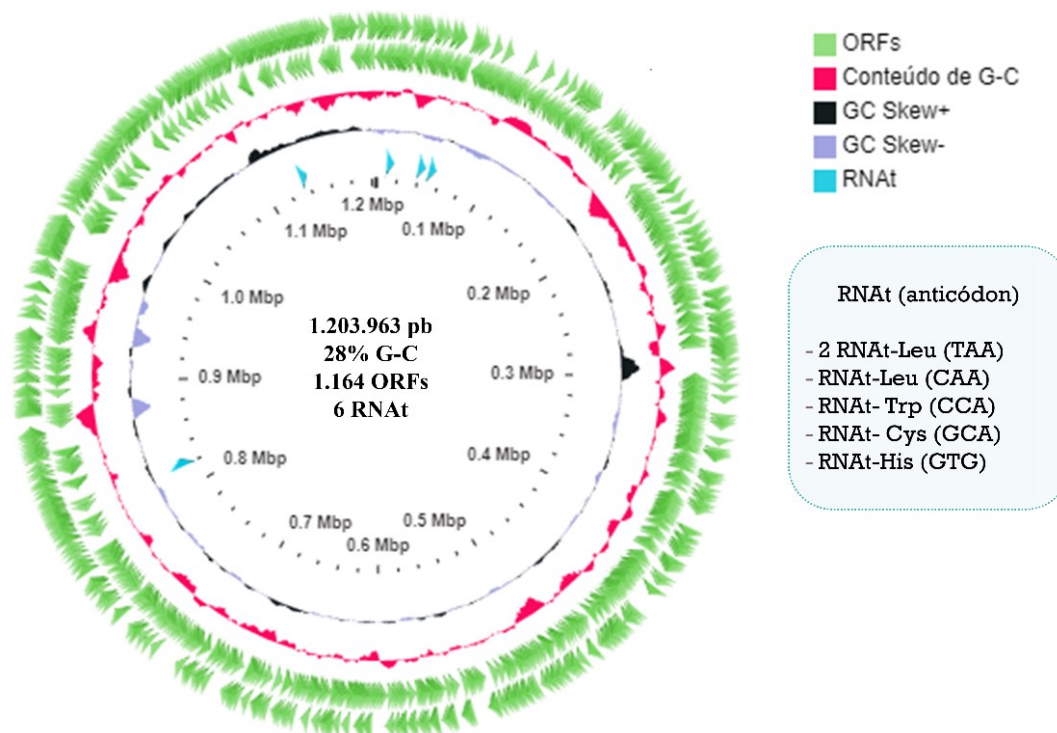


Figura 17 - Mapa do genoma do mimivírus *argentum*, indicando características como a distribuição das ORFs, conteúdo de G-C e RNAs transportadores na sequência. A caixa azul especifica os seis RNAt codificados pelo genoma do mimivírus *argentum*.

6.3.2 Filogenia do mimivírus *argentum*

Com o objetivo de analisar as relações filogenéticas do mimivírus *argentum* com os outros membros do grupo dos NCLDV, foi utilizada a sequência de aminoácidos da DNA polimerase

B (M731) predita para a construção da árvore. A análise da árvore filogenética obtida permitiu concluir que o mimivírus *argentum* está agrupado em um clado composto por vírus da linhagem A do gênero *Mimivirus*. Esse clado, por sua vez, está relacionado mais proximamente com outros dois clados que representam as linhagens B e C do gênero *Mimivirus*. (Figura 18). Em conjunto, esses três clados de mimivírus se agrupam com outros membros da família *Mimiviridae*, apresentando uma relação mais próxima com os tupanvírus do que com os klosneuvírus e o Cafeteria roenbergensis vírus.

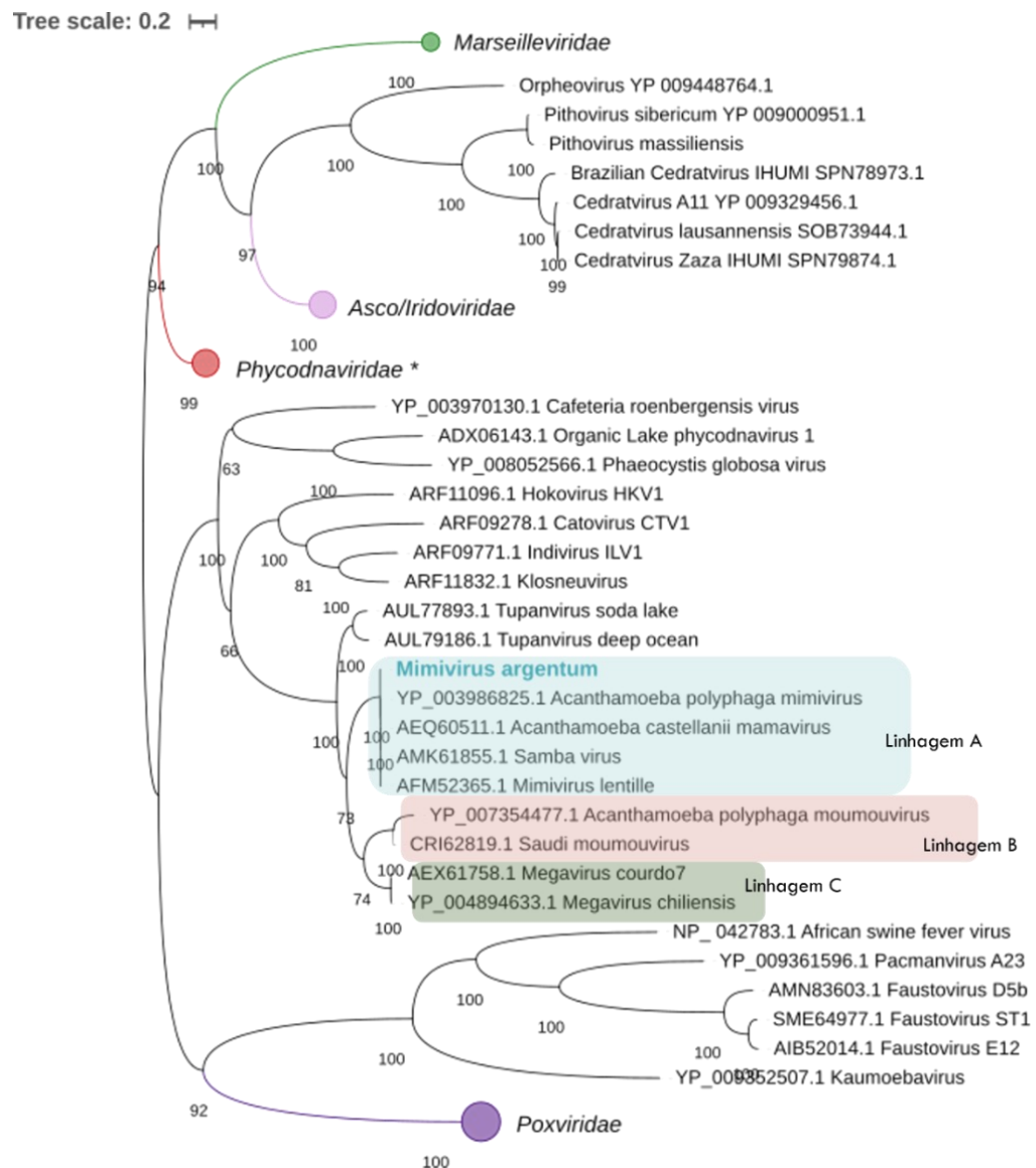


Figura 18 - Filogenia de *Nucleocytoviricota*, com base na sequência de aminoácidos da enzima DNA polimerase B, presente no mimivírus *argentum* (M731). O alinhamento foi realizado no MUSCLE, com parâmetros padrão. A árvore foi construída no IQtree, utilizando o método de máxima verossimilhança, com suporte estatístico estimado com base em 1000 replicatas (bootstrap). O melhor modelo, selecionado pelo IQtree (ModelFinder), para a essa árvore foi VT+F+R5. A árvore construída foi visualizada no software MEGAX e na plataforma online iTOL. * Nesta árvore, a junção do clado da família *Phycodnaviridae* inclui os pandoravírus e mollivírus.

De maneira geral, a filogenia do vírus isolado aqui apresentada reforça o que já era esperado, com base na microscopia eletrônica de transmissão e na caracterização genômica realizada. A presença de sequências de vírus que formam clados correspondentes a diferentes famílias do grupo dos NCLDV na árvore, bem como a relação dessas famílias com os mimivirus, indicam que o mimivírus *argentum* pode ser considerado um isolado de mimivírus da linhagem A, membro da família *Mimiviridae* e do filo *Nucleocytoviricota*.

6.4 Caracterização genômica do sputnik *argentum*

6.4.1 Anotação dos genes

O genoma do virófago sputnik *argentum* é formado por uma molécula de DNAfd circular, contendo 18.415 pb, com conteúdo de G-C de 26,96%. Em cada uma das extremidades da sequência de DNA do sputnik *argentum*, é possível observar uma repetição idêntica composta por 77 nucleotídeos. Utilizando o *GeneMarkS*, foram preditas 20 ORFs, com tamanhos que variam entre 330 e 2.340 nucleotídeos. As ORFs do virófago sputnik *argentum* são capazes de codificar proteínas cujos tamanhos variam de 109 a 779 aminoácidos. Os genes estão distribuídos em ambas as fitas do DNA, sendo que 15 estão localizados na fita senso e os outros 5 na fita antisenso. A análise das proteínas preditas, visando a busca por sequências similares no *BLASTp*, permitiu observar que todas elas são caracterizadas por possuir uma alta similaridade com os virófangos Sputnik 1 ou Sputnik 2, sempre apresentando 100% de identidade de aminoácidos e 100% de cobertura da sequência de entrada (*query*) com seus respectivos *best hits* (Tabela 6).

Tabela 6 – Resultados da análise das ORFs preditas para o virófago sputnik *argentum* no *BLASTp*.

ORF	Tamanho (nt)	Tamanho (aa)	Best hit <i>BLASTp</i>	Query cover	e-value	Identidade
1	2340	779	V13 [Sputnik virophage]	100%	0.0	100%
2	345	114	V14 [Sputnik virophage]	100%	1,00E-71	100%
3	330	109	V15 [Sputnik virophage]	100%	1,00E-70	100%
4	393	130	V16 [Sputnik virophage]	100%	2,00E-88	100%
5	564	187	Putative IS3 family transposase A protein [Sputnik virophage 2]	100%	5,00E-129	100%
6	1137	378	minor virion protein [Sputnik virophage 2]	100%	0.0	100%
7	1788	595	V20 [Sputnik virophage]	100%	0.0	100%
8	1329	442	V21 [Sputnik virophage]	100%	0.0	100%
9	435	144	V1 [Sputnik virophage]	100%	9,00E-94	100%

10	345	114	V2 [Sputnik virophage]	100%	1,00E-74	100%
11	738	245	V3 [Sputnik virophage]	100%	1,00E-178	100%
12	420	139	V4 [Sputnik virophage]	100%	2,00E-99	100%
13	360	119	V5 [Sputnik virophage]	100%	4,00E-75	100%
14	933	310	V6 [Sputnik virophage]	100%	0.0	100%
15	711	236	V7 [Sputnik virophage]	100%	8,00E-162	100%
16	555	184	V8 [Sputnik virophage]	100%	9,00E-126	100%
17	528	175	V9 [Sputnik virophage]	100%	2,00E-124	100%
18	681	226	V10 [Sputnik virophage]	100%	3,00E-163	100%
19	489	162	hypothetical protein Sputnik2_R11 [Sputnik virophage 2]	100%	8,00E-109	100.00%
20	459	152	V12 [Sputnik virophage]	100%	2,00E-99	100.00%

Notas: nt: nucleotídeos; aa: aminoácidos;

Após a análise inicial considerando os *best hits* obtidos no *BLASTp*, também foi realizada uma busca por domínios, repetições ou superfamílias de proteínas no *InterProScan*, bem como a revisão de trabalhos anteriores, envolvendo os virófagos sputnik. Isso foi feito visando a busca por possíveis funções ou domínios já preditos para essas proteínas, uma vez que as anotações no banco de dados do NCBI, especialmente para o virófago sputnik 1, não indicavam essas características. Com base nisso, foi possível identificar 12 proteínas, cujos homólogos já apresentavam domínios ou funções preditos, enquanto as outras 8 proteínas apresentavam funções ou domínios desconhecidos, sendo consideradas proteínas hipotéticas (Tabela 7). A busca no *InterProScan* não permitiu a predição de domínios, repetições ou superfamílias homólogas diferentes daquelas já propostas inicialmente. Assim, considerando apenas os *best hits* obtidos no *BLASTp* e a busca no *InterProScan*, seria possível concluir que 40% do genoma do sputnik argenteum seria composto por ORFans, que codificam “proteínas hipotéticas”.

Na tentativa de detectar sequências homólogas com uma maior sensibilidade, aprimorando a anotação e diminuindo, assim, a quantidade de “proteínas hipotéticas”, foi utilizado o *HHpred*. A partir dessa análise, dentre as 20 proteínas preditas para o sputnik argenteum, 4 apresentaram resultados diferentes do que já havia sido proposto anteriormente para os seus homólogos, sendo que 3 delas eram consideradas “proteínas hipotéticas”, sem domínio ou função preditos e detectados em bancos de dados (Tabela 7).

Tabela7 – Análise das proteínas do sputnik argentum no *HHpred*, em comparação com os domínios e funções preditos anteriormente para as suas sequências homólogas.

Proteína do sputnik argentum	Domínios/funções preditos anteriormente	Hit <i>HHpred</i>	Probabilidade.	<i>e-value</i>
S1	Primase–helicase	Primase; helicase {Nitratiruptor phage NrS-1}	99.95	8.5e-27
S2	Fita-Zinco	Zinc motif and AAA domain protein {Homo sapiens}	96.56	0.053
S3	Transmembrana	Sem hit	-	-
S4	Desconhecido	Sem hit	-	-
S5	transposase A família IS3	Transcriptional regulator {Corynebacterium glutamicum}	97.24	0.0016
S6	Proteína secundária do capsídeo	Minor virion protein {Sputnik virophage}	100	8.4e-112
S7	Proteína principal do capsídeo	Capsid protein V20 {Sputnik virophage}	100	6.8e-139
S8	Desconhecido	L544; Transferase {Acanthamoeba polyphaga mimivirus}	100	1.4e-40
S9	Desconhecido	C4-type zinc ribbon domain {Chlamydia trachomatis 434/Bu}	87.68	0.54
S10	Desconhecido	Sem hit	-	-
S11	ATPase superfamília FtsK-HerA	FtsK-HerA superfamily, P-loop ATPase {Sulfolobus turreted icosahedral virus 2}	99.53	2.9e-12
S12	Fita-Zinco	C2H2-type zinc finger {Drosophila melanogaster}	97.86	0.000017
S13	Desconhecido	Sem hit	-	-
S14	Repetição tripla-hélice de colágeno	Collagen alpha 2(I) chain {Rattus norvegicus}	97.66	0.0031
S15	Repetição tripla-hélice de colágeno	Collagen-like protein {Bacillus anthracis}	98.62	0.000032
S16	Proteína secundária do capsídeo	Sem hit	-	-
S17	Desconhecido	Cysteine proteinases superfamily {Mastadenovirus H2}	99.82	2.2e-19
S18	Integrase (tirosina recombinase)	Site-specific recombinase {Escherichia coli}	99.88	6.7e-21
S19	Desconhecido	Sem hit	-	-
S20	Desconhecido	Sem hit	-	-

Notas: Os hits do *HHpred* indicados em negrito, representam os resultados diferentes dos que foram preditos anteriormente para as proteínas homólogas às do sputnik argentum. APMV: Acanthamoeba polyphaga mimivirus.

Fonte: dados da coluna de “domínios/funções preditos anteriormente” obtidos de LA SCOLA *et al.*, 2008.

Dentre as proteínas preditas para o sputnik argentum, foi possível identificar o grupo de proteínas conservadas da família *Lavidaviridae*. Isso inclui a proteína principal do capsídeo (S7), duas proteínas secundárias do capsídeo (S6 e S16) e uma ATPase da superfamília FtsK-HerA (S11), que estão associadas com a morfogênese da partícula. Um fato interessante foi o hit do *HHpred* para a proteína S11, que indicou homologia com um vírus de arqueia da classe *Tectiliviricetes*. A cisteína protease (S17) também foi identificada, entretanto, essa função ainda

não havia sido anotada adequadamente para o virófago sputnik 1, sendo inferida após as análises no *HHpred* (Tabela 7). Apesar disso, a busca por homólogos das proteases de outros virófagos, na ocasião de estudos filogenéticos, já indicava a ocorrência dessa proteína nos virófagos sputnik, ainda que a anotação a caracterizasse como proteína hipotética.

Além disso, também foram identificadas três proteínas contendo domínio com dedo de zinco (S2, S9 e S12), sendo que uma delas (S9) era, anteriormente, considerada uma proteína hipotética. A presença desse tipo de domínio em proteínas, pode estar envolvida com a ligação ao DNA e com a regulação da transcrição. Outra proteína identificada no virófago sputnik argentum e que está presente em vários membros de *Lavidaviridae*, é uma proteína com domínio primase-helicase (S1) que pode estar relacionada com replicação de DNA do virófago.

Para além das proteínas conservadas em virófagos, também foram identificadas proteínas cuja ocorrência não é tão ampla dentro da família. A proteína S3 foi predita como uma proteína com domínio transmembrana, enquanto as proteínas S14 e S15 foram preditas como proteínas contendo repetição de tripla-hélice de colágeno. Essas proteínas tipo colágeno podem estar associadas com interações entre proteínas nas fábricas virais, como proposto para os virófagos do gênero *Sputnikvirus*. Outra proteína, para a qual uma possível função pode ser predita, foi a proteína S18, cujo homólogo é uma tirosina recombinase sítio-específica, cujas estruturas detectadas compõem proteínas de fago.

No que diz respeito à proteína S5 do sputnik argentum, as análises no *HHpred* entregaram um resultado sutilmente diferente do que havia sido predito anteriormente para os homólogos dessas proteínas (Tabela 7). Esses resultados mostram que a proteína S5 do sputnik argentum apresenta homologia com uma região contendo um domínio de hélice-volta-hélice (HTH, *Helix-turn-helix*) tipo TetR, presente em uma família de reguladores de transcrição de procariotos. Essa estrutura HTH está relacionada à interação de proteínas regulatórias com o DNA, sendo parte de um domínio de ligação ao DNA (NELSON; COX, 2019). Entretanto, a subunidade de transposase predita inicialmente também está associada com a função de ligação ao DNA e com uma origem procariótica. Por fim, além da cisteína protease (S17), outra proteína do sputnik argentum cuja possível função só pôde ser predita após as análises do *HHpred* foi a S8, que apresentou homologia com uma enzima transferase presente no APMV, com uma probabilidade de 100%. (Tabela 7)

De maneira geral, com a busca por sequências homólogas realizada por meio de diferentes metodologias e bancos de dados, foi possível predizer domínios ou funções para 15 das 20

proteínas do sputnik argentum, enquanto as outras 5 continuam sendo anotadas como proteínas hipotéticas (Figura 19 A). Isso indica que 25% do genoma do sputnik argentum é composto por ORFans, uma redução de 15 pontos percentuais em comparação à anotação realizada utilizando apenas os *best hits* do BLASTp. Quanto aos 75% das proteínas restantes, para as quais se propôs algum domínio ou função, foi possível observar que 20% do genoma é composto por genes que codificam proteínas com domínio de ligação ao DNA, bem como 25% de proteínas que estão relacionadas à morfogênese da partícula. Os genes para as proteínas contendo repetição de tripla-hélice de colágeno e a proteína com domínio primase-helicase, associada com a replicação do DNA representam proporções menores do genoma do sputnik argentum (10 e 5%, respectivamente) (Figura 19 B).

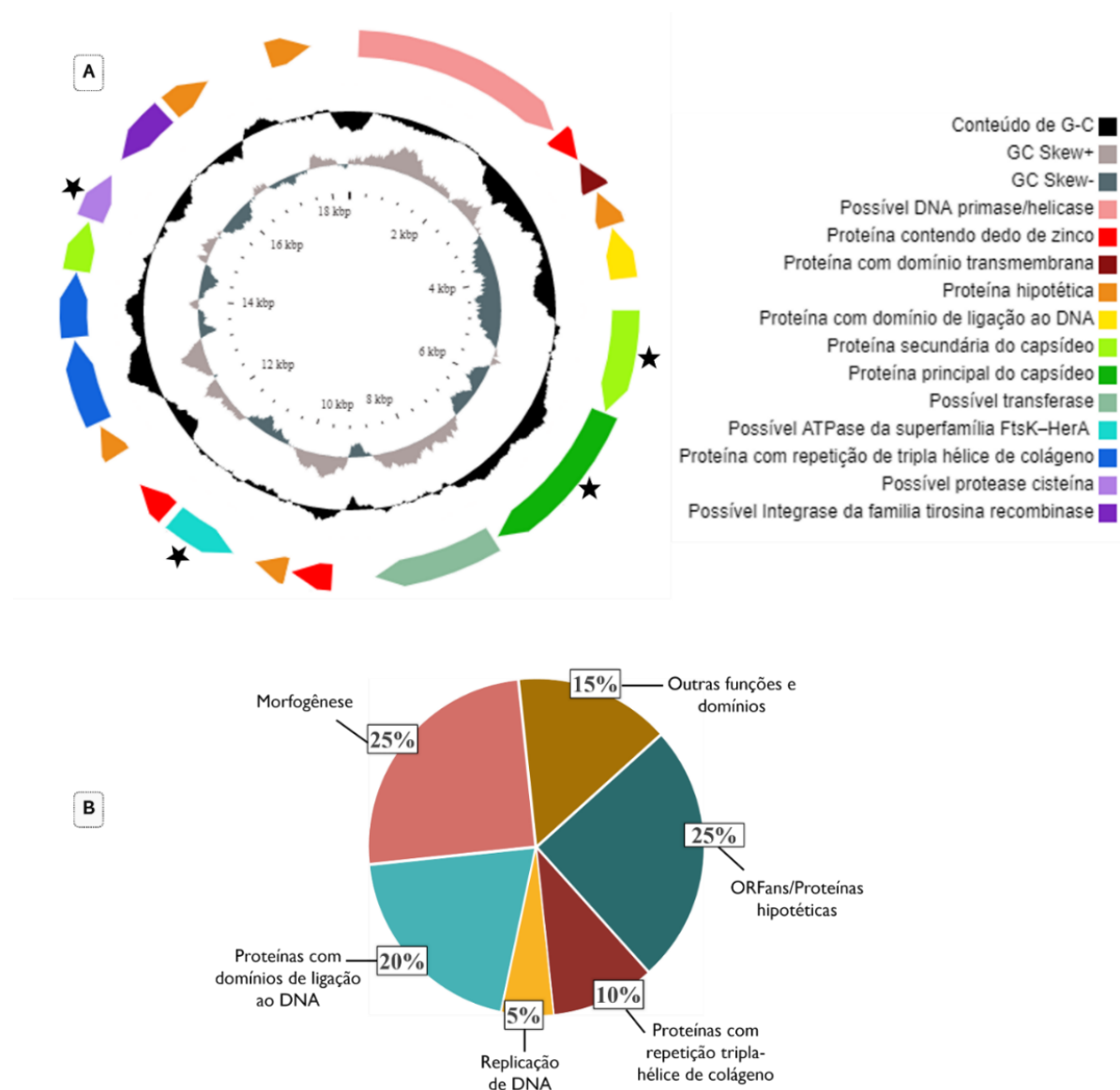


Figura 19 - Caracterização do genoma do virófago sputnik argentum. (A) Mapa do genoma do virófago sputnik argentum indicando, com cores, a predição de domínios, repetições ou funções para os produtos das 20

ORFs que compõem o seu genoma. As ORFs sinalizadas com uma estrela preta codificam as principais proteínas conservadas na família. (B) Gráfico que destaca a proporção, no genoma do sputnik argentum, de algumas das funções ou domínios/repetições presentes nas proteínas preditas.

6.4.2 Filogenia do sputnik argentum

As análises genômicas realizadas indicaram que o sputnik argentum é um virófago muito similar aos outros isolados do gênero *Sputnik*, como pode ser observado na tabela 8, que compara as características genômicas de cada isolado de virófago, pertencentes aos dois gêneros da família *Lavidaviridae*. Além disso, ao alinhar a sequência de DNA completa do sputnik argentum contra as sequências dos outros isolados de virófagos no *BLASTn*, foi possível reforçar a similaridade do virófago sputnik argentum com os virófagos sputnik 2 e rio negro, já que suas sequências apresentam 100% de identidade de nucleotídeos. Por outro lado, os virófagos mais divergentes em relação ao sputnik argentum são os virófagos zamilon e mavírus, como é esperado para virófagos do grupo dos sputnik. (Tabela 8).

Tabela 8 – Resultados do alinhamento (*BLASTn*) da sequência de nucleotídeos do sputnik argentum contra outros virófagos isolados e comparação de suas características genômicas.

Virófago	Genoma (pb)	Conteúdo de G-C (%)	Número de ORFs	Identidade de nt com o sputnik argentum (<i>BLASTn</i>)	Query cover	e-value
Sputnik argentum	18.415	26,96	20	-	-	-
Sputnik 1	18.343	27	20	99,98%	100%	0.0
Sputnik 2	18.338	27	20	100%	100%	0.0
Sputnik 3	18.338	27	20	99,96%	100%	0.0
Rio Negro	18.145	27	20	100%	98%	0.0
Guarani	18.967	26	22	97,95%	95%	0.0
Zamilon	17.276	30	20	84,38%	44%	0.0
Zamilon vitis	17.327	30	20	84,45%	39%	0.0
Mavírus	19.063	30	20	nenhuma similaridade significativa encontrada	-	-

Notas: pb: pares de bases, nt: nucleotídeos

Para reforçar esses dados, foram realizadas análises filogenéticas com base nas quatro proteínas do sputnik argentum envolvidas com a morfogênese das partículas e que são conservadas dentro da família *Lavidaviridae*. Essas proteínas foram escolhidas, pois são

sabidamente capazes de refletir as relações filogenéticas do grupo de maneira mais clara. Assim, foram construídas árvores filogenéticas individuais para a MCP (Figura 20 A), para uma das proteínas secundárias do capsídeo (Figura 20 B), para a ATPase (Figura 20 C) e para a cisteína protease (Figura 20 D) do sputnik argenteum.

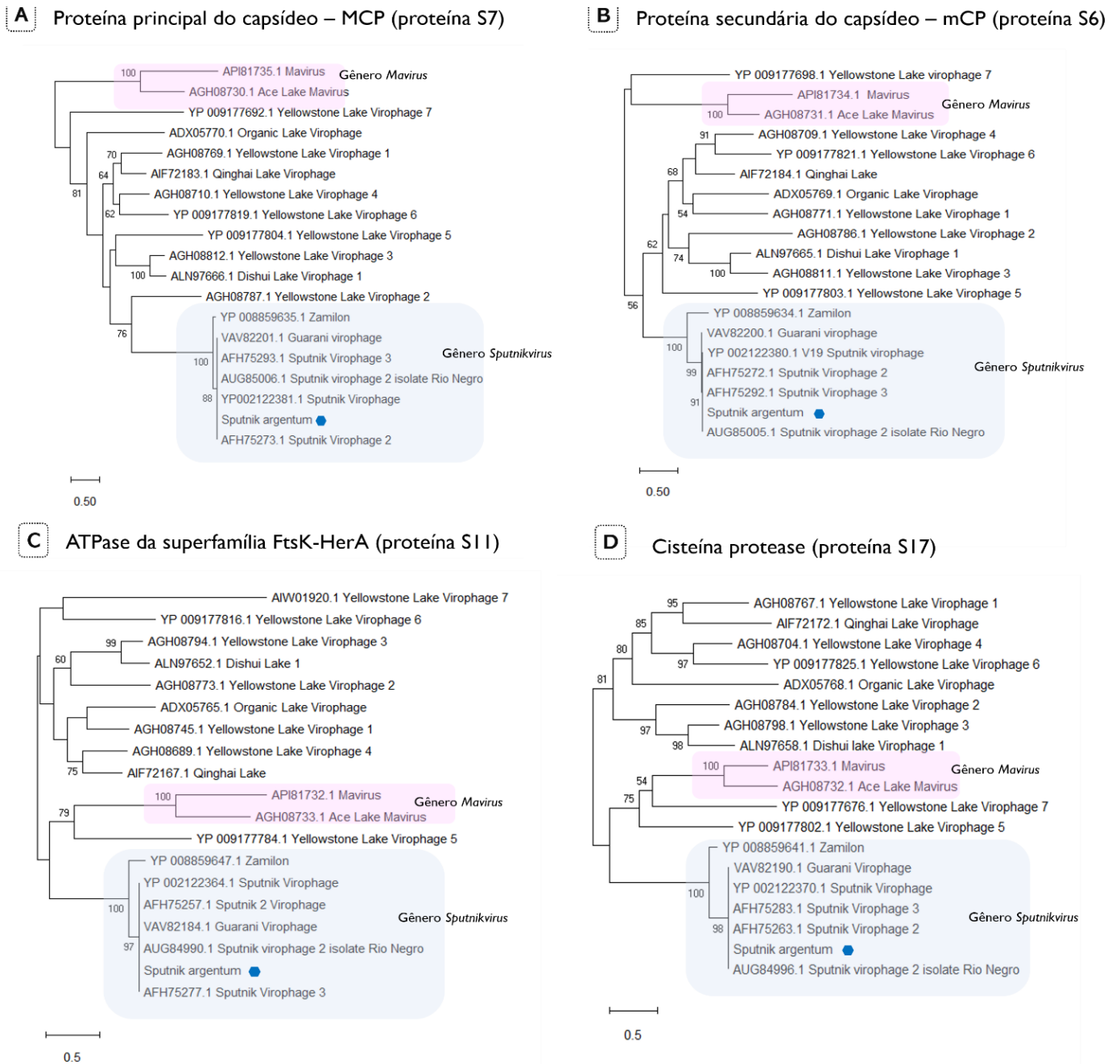


Figura 20 -Filogenia das proteínas S7, S6, S11 e S17 do virófago sputnik argenteum. (A-D) Árvores filogenéticas da proteína principal do capsídeo (S7), proteína secundária do capsídeo (S6), ATPase da superfamília FtsK-HerA (S11) e protease cisteína (S17), respectivamente. Os alinhamentos foram realizados no MUSCLE, com parâmetros padrão. As árvores foram construídas no IQtree, utilizando o método de máxima verossimilhança, com suporte estatístico estimado com base em 1000 replicatas (*bootstrap*). Os melhores modelos, selecionados pelo

IQtree (ModelFinder), para as árvores A, B, C e D foram rtREV+F+I+G4; VT+F+I+G4; rtREV+F+I+G4 e LG+F+R3, respectivamente. As árvores construídas foram visualizadas no software MEGAX.

Os dados obtidos após a elaboração de todas as quatro árvores levaram ao mesmo resultado, o qual indica que o isolado sputnik argentum se agrupa com um alto suporte estatístico (*bootstrap*) aos outros isolados de virófago sputnik em um clado que compõe o gênero *Sputnikvirus* (Figura 20). Além disso, a análise das árvores obtidas permite observar a separação dos cladogramas que representam as duas diferentes espécies do gênero *Sputnikvirus*, sendo uma delas representada pelo isolado zamilon e a outra representada por todos os isolados de sputnik. Também foi possível observar em todas as árvores a separação de outro clado, que contém o único isolado do gênero *Mavirus* e uma sequência relacionada, pertencente ao ace lake mavirus (ALM), cujo genoma foi detectado por metagenômica. A exemplo do ALM, as sequências que se agrupam fora dos cladogramas que compõem os dois gêneros de virófagos são todas pertencentes a virófagos detectados por metagenômica. Dentre essas sequências, algumas de suas relações filogenéticas também se mantêm em todas as árvores obtidas, como se observa para o clado composto pelos virófagos dishui lake e yellowstone lake 3 (Figura 20). Em suma, essas análises filogenéticas permitem concluir que o virófago sputnik argentum pode ser considerado um novo isolado de virófago sputnik 2, membro da família *Lavidaviridae* e do gênero *Sputnikvirus*.

6.5 Mimivirus argentum e seu mobiloma

6.5.1 Detecção do provirófago do sputnik argentum no genoma do mimivírus argentum

Os mimivírus são vírus singulares, pois, dentre tantos outros motivos, são capazes de carregar seus próprios elementos genéticos móveis no genoma, o que indica que eles apresentam um mobiloma próprio. As inferências realizadas na ocasião do isolamento do mimivírus argentum, bem como a anotação de seus genes, descrita neste trabalho, sugeriram que o genoma sputnik argentum estivesse integrado na sequência do vírus gigante hospedeira.

Ao analisar a sequência de DNA do mimivírus argentum, obtida após o sequenciamento, foi realizado um alinhamento entre a sequência do virófago sputnik argentum e a do mimivírus argentum na plataforma *BLASTn* (*expect threshold*: 10^{-3}). Os resultados mostraram um alinhamento com 100% de identidade de nucleotídeos e de cobertura, o que significa que 100%

da sequência de bases presentes no virófago foram alinhados com uma região do genoma do vírus localizada entre os nucleotídeos 938.070 e 956.484 (Figura 21).

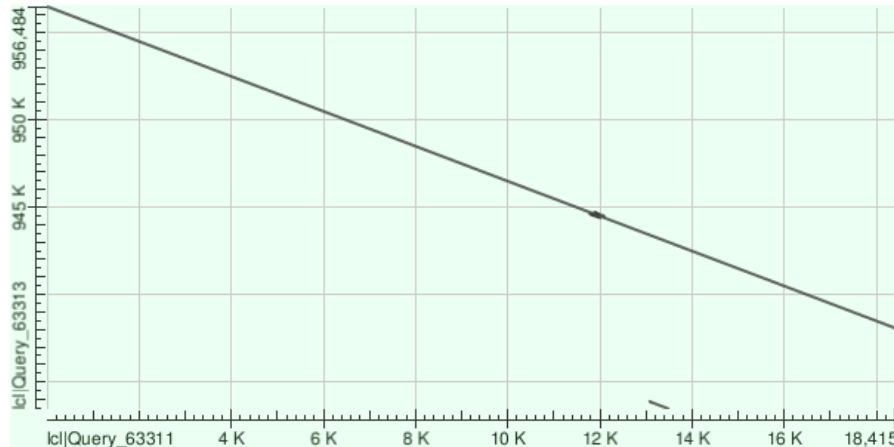


Figura 21 - Gráfico de pontos (*Dot plot*) gerado após alinhamento das sequências do virófago sputnik argentum e do mimivírus argentum, utilizando o *BLASTn*. O eixo X representa a posição dos nucleotídeos do sputnik argentum e o eixo Y representa a posição dos nucleotídeos do mimivírus argentum. A reta representa a correspondência entre a posição dos nucleotídeos de virófago com as posições de nucleotídeos do vírus gigante (entre 938.070 e 956.484).

Por outro lado, ao utilizar o mesmo método para analisar a sequência do virófago sputnik 2 contra o genoma do lentillevírus, o primeiro sistema para o qual foi descrito esse tipo de integração, não foi possível detectar uma região de inserção de toda a sequência assim como observado para o mimivírus argentum e sputnik argentum. O genoma do lentillevírus está depositado nos bancos de dados (*GenBank*) em 10 diferentes *contigs* e ao realizar um alinhamento do virófago sputnik 2 contra cada um desses *contigs*, uma região de integração completa do virófago não foi detectada. Apesar disso, o alinhamento da sequência do sputnik 2 com o contig 5, revelou uma pequena região, de 423 nucleotídeos, que apresentou 91,49% de identidade de nucleotídeos (Tabela 9). Essa região do genoma do virófago sputnik 2 que se alinha com o lentillevírus corresponde a genes para proteínas com repetição de tripla-hélice de colágeno em ambos os vírus.

Paralelamente, na figura 21, é possível observar, além da reta que representa o alinhamento de toda a sequência do sputnik argentum com o genoma do mimivírus argentum, uma pequena região alinhada. Essa pequena região alinhada corresponde às posições 13.083-13.504 do virófago e 933.444-933.866 do vírus gigante, sendo que tanto no sputnik argentum, quanto no mimivírus argentum essas posições também correspondem a proteínas com as repetições de

colágeno. Os alinhamentos detectados nessa pequena região, tanto no sistema sputnik2-lentilevírus quanto no sistema sputnik argenteum-mimivírus argenteum, são indicativos de que essas proteínas com repetições de colágeno possuem homólogos em virófagos e em vírus gigantes.

Tabela 9 – Resultados do alinhamento (BLASTn) entre a sequência do virófago sputnik 2 contra o genoma do lentilevírus, fragmentado em 10 diferentes contigs.

<i>Contig</i>	<i>Número de acesso</i>	<i>Query cover</i>	<i>e-value</i>	<i>Identidade</i>
1	AFYC01000001.1	Nenhuma similaridade significativa encontrada		
2	AFYC01000002.1	Nenhuma similaridade significativa encontrada		
3	AFYC01000003.1	Nenhuma similaridade significativa encontrada		
4	AFYC01000004.1	Nenhuma similaridade significativa encontrada		
5	AFYC01000005.1	2%	2,00E-165	91.49%
6	AFYC01000006.1	Nenhuma similaridade significativa encontrada		
7	AFYC01000007.1	Nenhuma similaridade significativa encontrada		
8	AFYC01000008.1	Nenhuma similaridade significativa encontrada		
9	AFYC01000009.1	Nenhuma similaridade significativa encontrada		
10	AFYC01000010.1	Nenhuma similaridade significativa encontrada		

Como apontado anteriormente, durante a anotação do genoma do mimivírus argenteum, foi possível detectar genes para várias proteínas altamente similares às de virófagos inseridas no genoma. As análises posteriores desses genes revelaram que todas as 20 ORFs do virófago sputnik argenteum foram anotadas junto ao genoma do mimivírus argenteum, em uma posição que corresponde à região de integração proposta no alinhamento do *BLASTn*. Em conjunto esses dados sugerem que essa seja a primeira descrição da integração completa de um provirófago no genoma do seu vírus gigante hospedeiro.

O próximo passo foi, então, analisar quais proteínas fazem parte da região vizinha desse possível ponto de integração do virófago. Essa análise permitiu observar que em ambas as extremidades da região de possível inserção do genoma do virófago, estão presentes alguns genes para proteínas com repetição de tripla-hélice de colágeno (Figura 22). A busca por genes para essas proteínas em outras regiões do genoma do mimivírus argenteum revelou que eles também estão presentes em conjunto, em outras três regiões do genoma, mais afastadas da região de integração do virófago. (Figura 22, em laranja)

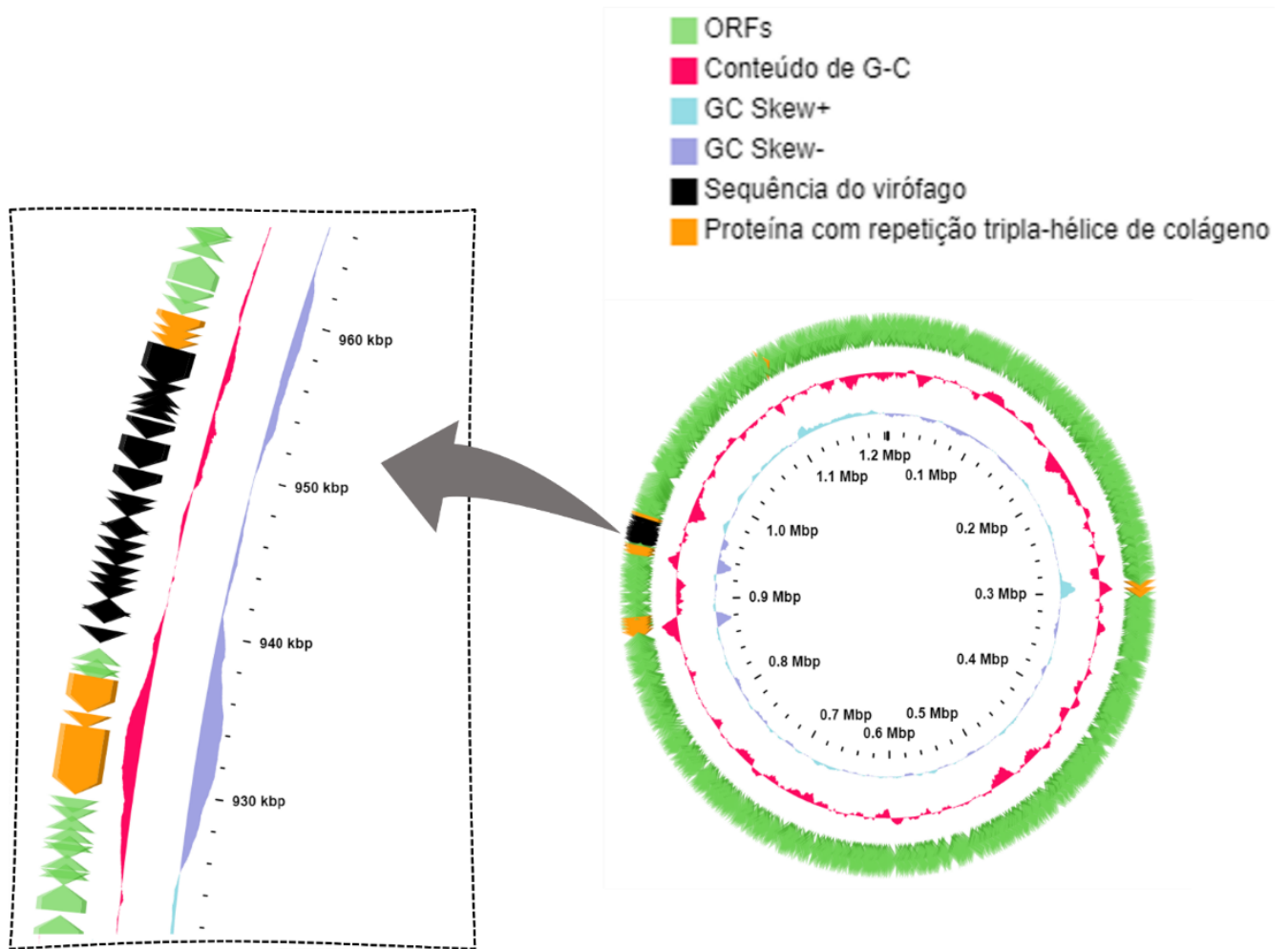


Figura 22 – Mapa do genoma do mimivírus argentum, destacando a possível região de integração da sequência do virófago sputnik argentum. Em destaque, é possível observar as 20 ORFs do virófago (preto), presentes em uma região (938.070 a 956.484) do genoma do mimivírus argentum que é cercada nas duas extremidades por proteínas com repetição de tripla-hélice de colágeno (laranja). Essas proteínas com repetição de colágeno também podem ser observadas em outras três regiões do genoma do mimivírus argentum.

6.5.2 Análise comparativa das proteínas contendo repetição de tripla-hélice de colágeno presentes em virófagos e vírus gigantes.

Considerando a presença desses genes para proteínas contendo repetição de tripla hélice de colágeno na região onde o sputnik argentum está, possivelmente, integrado, foi realizada uma análise quantitativa da presença desse tipo de gene em outras sequências de mimivírus. Foram escolhidas a sequência do mamavírus, por ser mais similar à do mimivírus argentum, a sequência do lentillevírus por ser o vírus para o qual foi descrita pela primeira vez a presença de provirófagos e a sequência do APMV, por ser o vírus protótipo do grupo. Além disso, as sequências do *Acanthamoeba polyphaga* moudouvírus e do megavírus chilensis também foram

utilizadas para a comparação, já que eles são os principais representantes das linhagens B e C de mimivírus, respectivamente.

Essa análise revelou a presença de 14 genes para proteínas contendo repetição de colágeno distribuídos pelo genoma do mimivírus *argentum*, sendo que 6 deles estão localizados na região de vizinhança da possível integração do virófago. No caso dos outros vírus, foram encontradas 11, 9 e 8 proteínas contendo repetição de colágeno em mamavírus, lentillevírus e APMV, respectivamente. Já para o *moumouvírus*, foram encontradas 7 dessas proteínas, enquanto na sequência do *megavírus chilensis* foram encontradas 13 (Figura 23). Dessa forma, é possível perceber que o *mimivírus argentum* apresenta uma maior quantidade de proteínas contendo repetição de tripla-hélice de colágeno, principalmente, em comparação aos outros mimivírus de linhagem A e ao *moumouvírus*. O *megavírus chilensis* apresenta uma quantidade parecida dessas proteínas, entretanto, o seu genoma também é maior em comparação aos outros mimivírus. Esses dados podem ser um indicativo de que proteínas com esse tipo de repetição estejam relacionadas com a integração de provirófagos nos genomas de mimivírus.

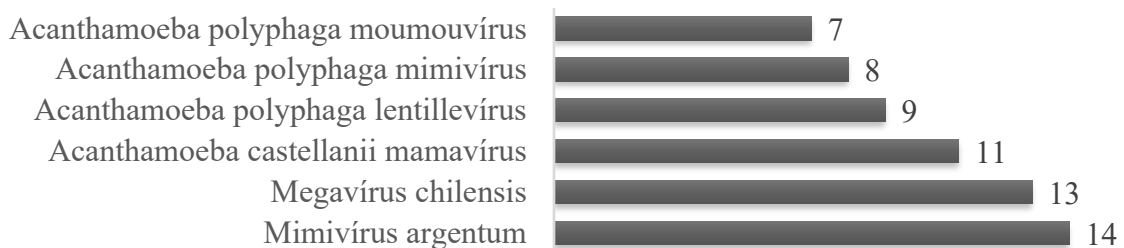


Figura 23 - Número de proteínas contendo repetição de tripla hélice de colágeno em diferentes mimivírus.

As proteínas contendo as repetições de tripla hélice de colágeno também são codificadas pelo genoma dos virófagos. Tendo isso em vista, também foi realizada uma análise quantitativa e comparativa da presença dessas proteínas em virófagos do gênero *Sputnikvirus*. Dessa forma, foi possível observar que as proteínas contendo repetição de colágeno estão presentes em todos os virófagos do gênero *Sputnikvirus* analisados e que os genes que as codificam aparecem sempre em pares, mantendo a sintenia em todos os genomas (Figura 24). É interessante ressaltar que, em comparação aos virófagos sputnik 1, 2, 3 e rio negro, os genomas de outros virófagos como o sputnik *argentum*, o guarani e os dois virófagos zamilon aparentam não possuir a mesma sintenia de genes. Entretanto, é possível observar que a ordem dos genes se mantém e isso se deve ao fato de os genomas desses virófagos serem moléculas de DNA com topologia circular, apesar de estarem representados de maneira linear na Figura 24.

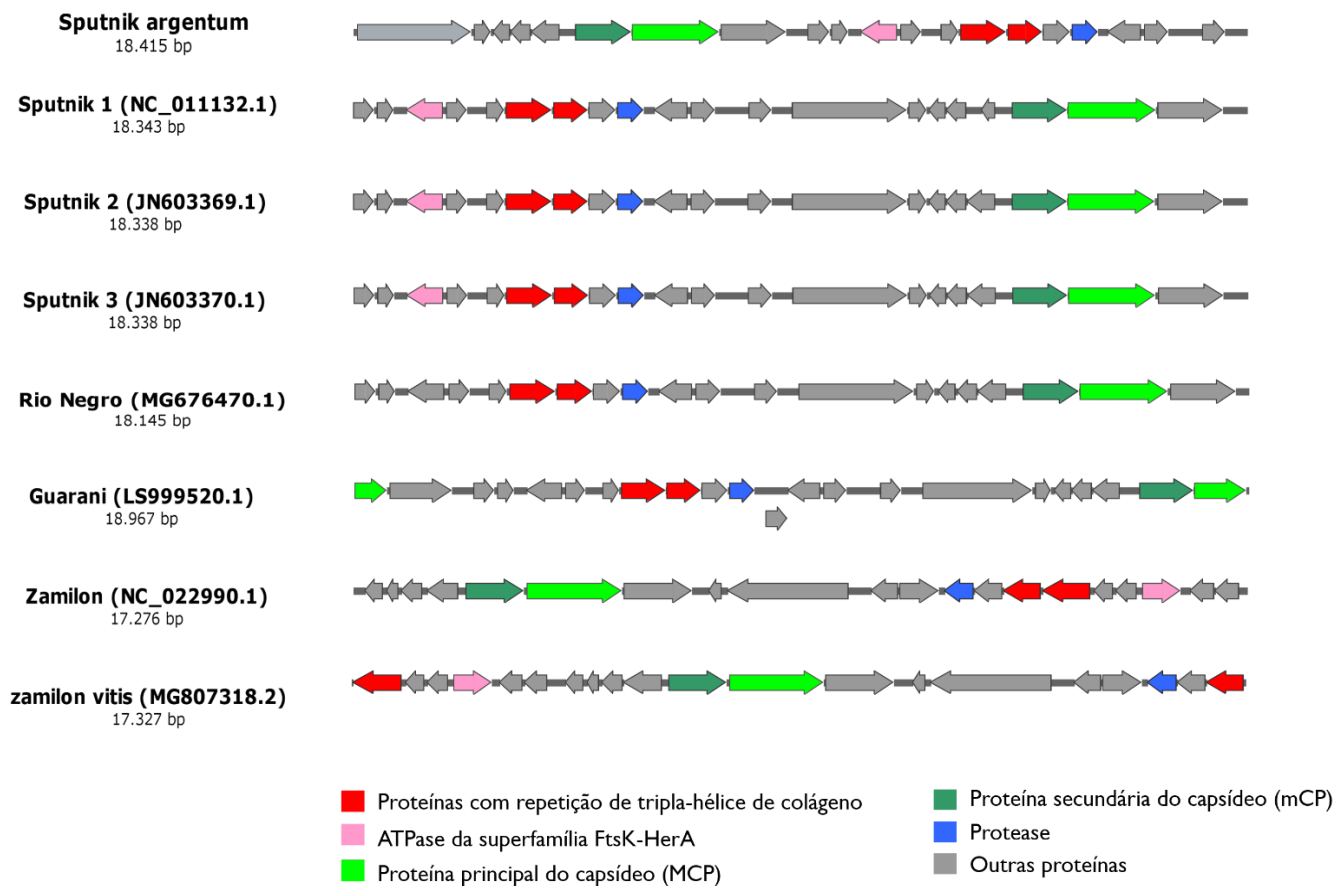


Figura 24 – Análise de sintenia dos genes para as proteínas contendo repetição de tripla hélice de colágeno presentes em diferentes virófagos do gênero *Sputnikvirus*. É possível observar os genes para as proteínas com repetição de colágeno ocorrendo em pares (setas vermelhas). Os genes para proteínas conservadas dentro os virófagos também estão indicados com outras cores (legenda). Obs1.: As setas direcionadas para a direita indicam genes posicionados no sentido 5'-3' do DNA, enquanto as setas direcionadas para a esquerda indicam os genes posicionados no sentido 3'-5'. Obs2: É válido ressaltar que os genomas dos virófagos aqui representados têm topologia circular, embora estejam representados de forma linear.

Adicionalmente, foi determinada a porcentagem de identidade de nucleotídeos, ao alinhar no *BLASTn*, o par de genes para as proteínas com repetição de colágeno dos virófagos analisados contra o genoma do virófago sputnik argentum. Essa análise indicou que os dois genes, presentes no virófago sputnik argentum, são muito similares aos dos outros virófagos analisados. Ambos os genes codificados pelos virófagos 1, 2, 3 e rio negro apresentaram 100% de identidade de nucleotídeos quando alinhados contra o genoma do sputnik argentum (Tabela 10). Os genes codificados pelo virófago guarani também são bastante similares ao sputnik argentum, sendo que um deles apresentou 100% de identidade e o outro 99,68%. Já em comparação com os virófagos zamilon, a similaridade de genes com o sputnik argentum se mostra mais reduzida. (Tabela 10). Em conjunto, esses dados sugerem que as proteínas com

repetição de colágeno podem ser consideradas conservadas dentre os virófangos do gênero *Sputnikvirus*.

Tabela 10 – Resultados da comparação, realizada no BLASTn, entre as proteínas contendo repetição de tripla hélice de colágeno de virófangos do gênero *Sputnikvirus*.

Virófago	Proteína	Nº de acesso (GenBank)	Tamanho (aa)	BLASTn contra o genoma do <i>Sputnik argentum</i>		
				Query Cover	e-value	Identidade
S. argentum	S14	-	310	-	-	-
S. argentum	S15	-	236	-	-	-
Sputnik 1	V6	ACF16990.1	310	100%	0.0	100.00%
Sputnik 1	V7	ACF16991.1	236	100%	0.0	100.00%
Sputnik 2	R6	AFH75260.1	310	100%	0.0	100.00%
Sputnik 2	R7	AFH75261.1	236	100%	0.0	100.00%
Sputnik 3	R6	AFH75280.1	310	100%	0.0	100.00%
Sputnik 3	R7	AFH75281.1	236	100%	0.0	100.00%
Rio Negro	6	AUG84993.1	310	100%	0.0	100.00%
Rio Negro	7	AUG84994.1	236	100%	0.0	100.00%
Guarani	GUARANI_8	VAV82187.1	310	100%	0.0	99.68%
Guarani	GUARANI_9	VAV82188.1	236	100%	0.0	100.00%
Zamilon	X812_gp14	YP_008859643.1	241	100%	1,00E-144	79.64%
Zamilon	X812_gp15	YP_008859644.1	305	Nenhuma similaridade significativa encontrada		
Zamilon vitis	za3_20	AVL93360.1	241	100%	1,00E-149	80.03%
Zamilon vitis	za3_1	AVL93341.2	308	Nenhuma similaridade significativa encontrada		

Notas: aa: aminoácidos.

Com base nas análises comparativas realizadas acerca das proteínas com repetição de colágeno, a proteína S14 do virófago sputnik argentum foi utilizada para a construção de árvores filogenéticas. Dessa forma, foi possível concluir que as proteínas com repetição de colágeno, de fato, apresentam homólogos presentes nos virófangos e em membros do filo *Nucleocytoviricota*, como é o caso dos pithovirus, cedratvirus, pandoravirus e das famílias *Phycodnaviridae* e *Mimiviridae*. A árvore filogenética da proteína S14 indica que essa proteína é bastante conservada dentre os virófangos do gênero *Sputnikvirus*, sendo mais divergente na linhagem dos virófangos zamilon (Figura 25). Isso reflete a filogenia de outras proteínas conservadas em virófangos, que sustentam a taxonomia do grupo e permitem a divisão do gênero *Sputnikvirus* em duas espécies distintas no ICTV.

No que diz respeito ao clado que compõe família *Mimiviridae*, a filogenia também reforça a divergência de algumas linhagens já conhecidas anteriormente, como as três linhagens

(A, B e C) do gênero *Mimivirus* e a linhagem dos tupanvírus, proposta para representar um novo gênero para a família. Além disso, é possível observar que o clado formado pelos virófagos se origina a partir do mesmo ancestral que deu origem ao clado dos mimivírus, compartilhando um ancestral mais próximo com os mimivírus da linhagem A, do que com outros membros da família *Mimiviridae* (Figura 25). Por fim, as proteínas com repetição de colágeno de virófagos parecem ter uma relação evolutiva mais próxima com os mimivírus do que com os outros membros de *Nucleocytoviricota* representados na árvore, os pitho/cedratvírus, os pandoravírus e os *Phycodnaviridae* (Figura 25).

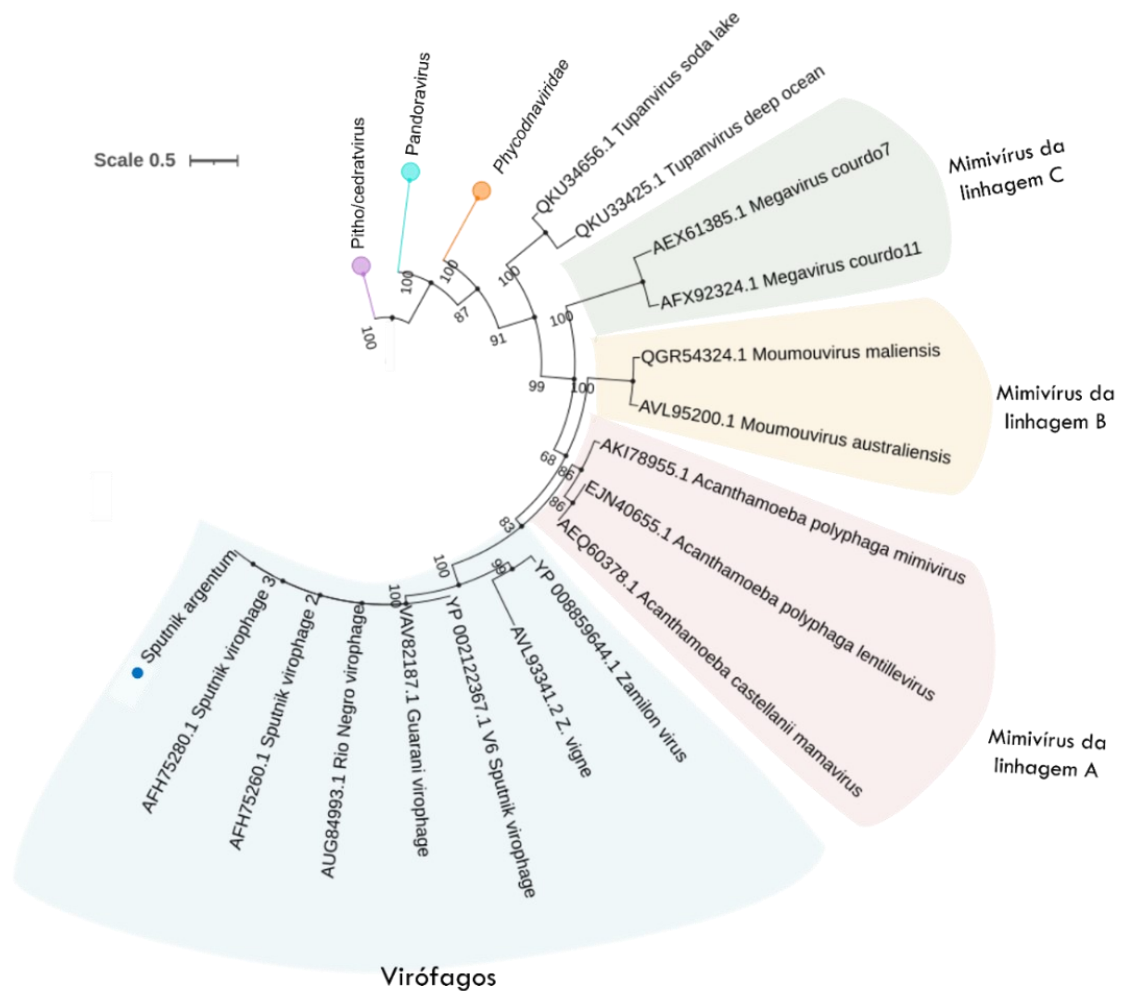


Figura 25 - Filogenia de uma das proteínas contendo repetição de tripla hélice de colágeno (S14) do virófago sputnik argenteum. O alinhamento foi realizado no MUSCLE, com parâmetros padrão. A árvore foi construída no IQtree, utilizando o método de máxima verossimilhança, com suporte estatístico estimado com base em 1000 replicatas (*bootstrap*). O melhor modelo, selecionado pelo IQtree (ModelFinder), para a essa árvore foi PMB+F+I+G4. A árvore construída foi visualizada na plataforma online iTOL.

6.5.3 Caracterização do transpoviron do mimivírus *argentum* e comparação com os outros transpovirons descritos.

6.5.3.1 Genoma do transpoviron do mimivírus *argentum*

Assim como os provirófagos, os transpovirons também fazem parte do mobiloma dos mimivírus. Como citado anteriormente, nesse trabalho foi detectada uma sequência de transpoviron associada ao mimivírus *argentum*. O genoma do transpoviron do mimivírus *argentum* (Tvmarg) é formado por uma molécula de DNAfd, que contém 6.624 pb e conteúdo de G-C de 24,86%. A predição de genes e de proteínas feita pelo *GeneMarkS* revelou que o Tvmarg apresenta 6 ORFs, com tamanhos que variam de 270 a 3.000 nucleotídeos. Já as proteínas codificadas por essas ORFs possuem entre 89 e 999 aminoácidos. Os 6 genes do Tvmarg estão distribuídos em ambas as fitas do DNA, sendo que 2 deles estão posicionados na fita senso e os outros 4 na fita antisenso.

A análise das ORFs preditas para o Tvmarg, realizada no *BLASTp*, revelou 5 proteínas que apresentam uma alta similaridade com sequências do transpoviron do lentillevírus (ou mimivirus lentille). A proteína restante apresentou similaridade com uma sequência do transpoviron do mamavírus. À exemplo do virófago sputnik *argentum*, para todas as proteínas, os valores de identidade de aminoácidos e de cobertura da sequência de entrada (*query*) foram de 100% (Tabela 11). Dentre os seis *best hits* do *BLASTp* obtidos para as proteínas do Tvmarg, 4 deles indicaram proteínas hipotéticas como similares, o que representa dois terços de suas ORFs. As outras duas sequências, apresentaram similaridade com a proteína com domínio de helicase e integrase do transpoviron do mamavírus e com uma proteína contendo dedo de zinco tipo C2H2 do transpoviron do lentillevírus (Tabela 11).

Tabela 11 – Análise das proteínas do transpoviron do mimivírus *argentum* utilizando o *BLASTp*

ORF	Tamanho (nt)	Tamanho (aa)	Best hit <i>BLASTp</i>	Query cover	e-value	Identidade
1	606	201	hypothetical protein tv_R1 [Mimivirus lentille]	100%	4,00E-142	100%
2	423	140	hypothetical protein tv_L2 [Mimivirus lentille]	100%	4,00E-92	100%
3	411	136	hypothetical protein tv_L3 [Mimivirus lentille]	100%	1,00E-87	100%
4	339	112	hypothetical protein tv_L4 [Mimivirus lentille]	100%	8,00E-73	100%
5	3000	999	integrase zinc-binding and helicase domain-containing protein [Acanthamoeba castellanii mamavirus]	100%	0.0	100%
6	270	89	zinc finger C2H2-type domain-containing protein [Mimivirus lentille]	100%	8,00E-57	100%

Notas: nt: nucleotídeos; aa: aminoácidos

Levando em conta a alta quantidade de proteínas hipotéticas que são consideradas como *best hits* (*BLASTp*) das proteínas do Tvmarg, foram feitas análises no *InterProScan*, no *HHpred* e em trabalhos anteriores. O objetivo das análises foi preencher essas lacunas, assim como o realizado para as sequências do mimivírus argenteum e do virófago sputnik argenteum. A partir dessa metodologia, foi possível identificar possíveis domínios para duas proteínas do sputnik argenteum, que no *BLASTp* eram consideradas hipotéticas. Os dados do *InterProScan* e do *HHpred* indicaram que a proteína T2 apresenta um domínio de hélice transmembrana (Tmhelix) e que a proteína T4 é homóloga a uma sequência possivelmente de uma transposase contendo o motivo HTH de ligação ao DNA, que, nos outros transpovirons descritos, foi proposta para estar associada tanto com a regulação da transcrição, quanto com a transposição (Tabela 12). As outras duas proteínas (T5 e T6) que apresentam algum tipo de função anotada são consideradas altamente conservadas dentre os transpovirons. As análises do *InterProScan* e *HHpred* reforçaram os dados do *BLASTp*, indicando a existência de um domínio de helicase na proteína T5 e de um dedo de zinco na proteína T6 (Tabela 12).

Tabela 12 – Dados da busca por domínios e sequências homólogas para as proteínas do transpoviron do mimivírus argenteum realizadas no *InterProScan* e no *HHpred*.

Proteína	Domínios InterProScan	Hit Hhpred	Prob.	e-value
T1	Sem hit	Sem hit	-	-
T2	Tmhelix, transmembrane region	Transmembrane protein: alpha-helical {Escherichia coli}	91.44	1
T3	Coil	Sem hit	-	-
T4	Domain 2r0qC02	Transposase {Corynebacterium glutamicum}	98.52	9.4e-7
T5	Viral helicase	Exodeoxyribonuclease v alpha chain; recombination, helicase, nuclease {Escherichia coli}	100	1.9e-31
T6	Sem hit	Zinc finger homeobox protein {Homo sapiens}	95.21	0.055

Notas: As linhas destacadas em negrito indicam os resultados que se diferenciam da anotação realizada somente utilizando o *BLASTp*. Prob.: probabilidade.

Em conjunto, os resultados da anotação dos genes do Tvmarg utilizando diferentes metodologias revela que ele apresenta 4 proteínas com possíveis funções ou domínios preditos, enquanto outras duas, as proteínas T1 e T3, continuaram sendo anotadas como proteínas hipotéticas, por não apresentarem nenhuma sequência homóloga nos bancos de dados que permitisse fazer essas inferências (Figura 26A). Dessa forma, o conteúdo de ORFans que

codificam as proteínas hipotéticas do Tvmarg foi reduzido pela metade, já que deixou de representar dois terços (4/6) para compor apenas um terço (2/6) do seu genoma.

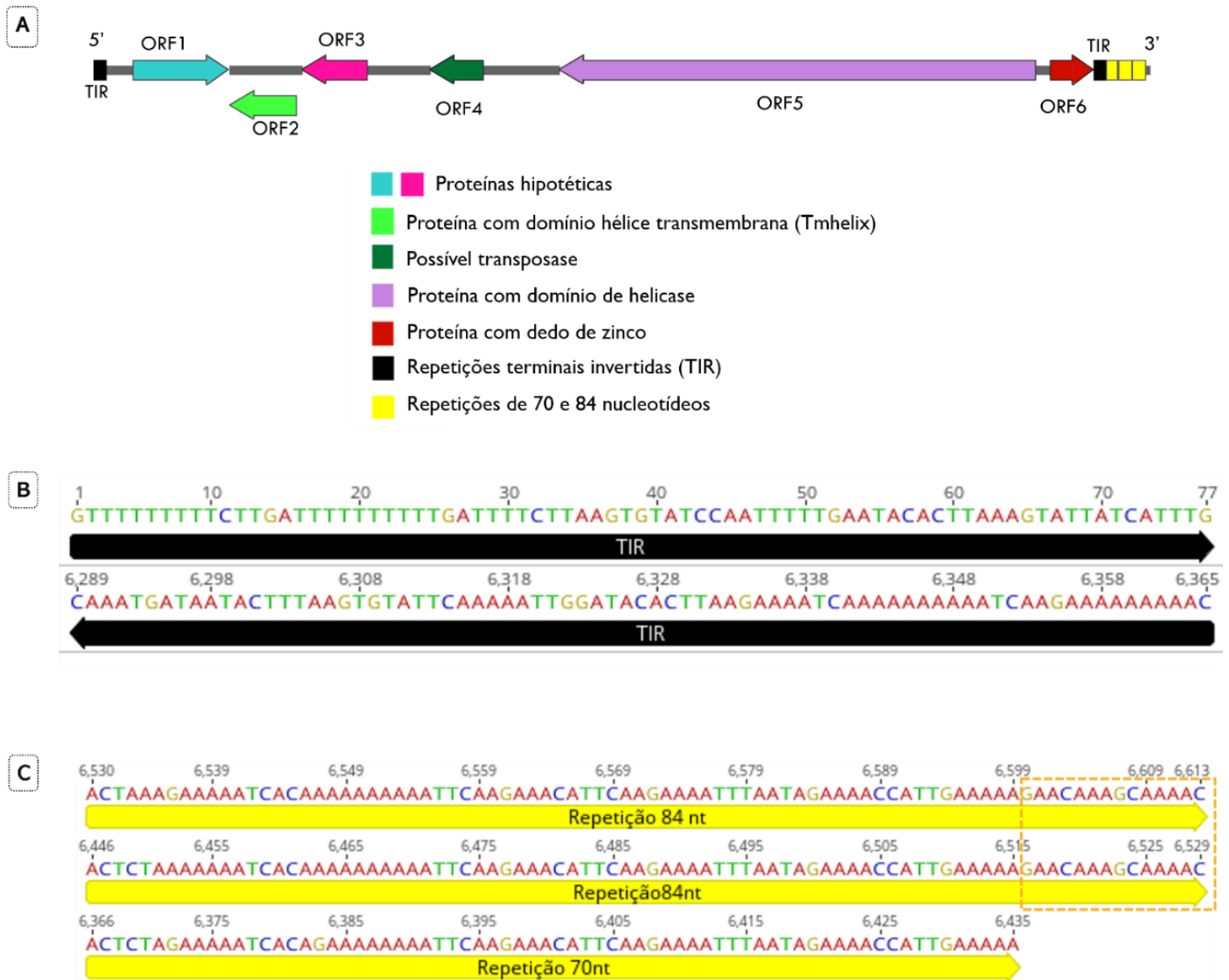


Figura 26 – Caracterização genômica do transpoviron do mimivírus *argentum* (Tvmarg). (A) Mapa do genoma do Tvmarg ilustrando a presença das 6 ORFs e suas respectivas anotações indicadas pelas cores na legenda. A presença das regiões terminais invertidas (TIRs) e de um conjunto de 3 repetições de 70 e 84 nucleotídeos (nt) também pode ser observada. (B) Figura mostrando a combinação dos nucleotídeos (100%) e a posição das TIRs no genoma do Tvmarg. (C) Figura exibindo a combinação de nucleotídeos das repetições de 84 nt com a repetição de 70 nt e suas respectivas posições no genoma do Tvmarg. É possível observar, em destaque, uma região de 14 nt presente exclusivamente nas repetições de 84 nt. As figuras foram elaboradas por meio dos softwares *SnapGene* e *GeneiousPrime*. Obs.: As setas direcionadas para a direita indicam genes posicionados no sentido 5'-3' do DNA, enquanto as setas direcionadas para a esquerda indicam os genes posicionados no sentido 3'-5'.

À exemplo dos outros transpovirons, o Tvmarg também apresenta repetições terminais invertidas em suas extremidades. Essas repetições presentes nas duas extremidades do genoma do Tvmarg são compostas por 77 nucleotídeos e apresentam 100% de combinação (*match*) entre si. A TIR da extremidade 5', que fica antes da ORF 1, está localizada entre as posições 1 e 77 do genoma, enquanto a TIR da extremidade 3' que fica imediatamente após a ORF6, está localizada entre as posições 6.289 e 6.365 do genoma (Figura 26 A e B). Se comparado aos outros transpovirons, o Tvmarg apresenta TIRs muito menores, uma vez que é descrito que eles possuam TIRs maiores do que 300 nucleotídeos, chegando a 520 nucleotídeos em transpovirons do clado A. Além das repetições terminais invertidas, o Tvmarg também possui um conjunto de três repetições diretas localizado após a TIR da extremidade 3'. Uma dessas repetições contém 70 nucleotídeos, enquanto as outras duas contêm 84 nucleotídeos. A repetição de 70 nucleotídeos faz uma combinação quase perfeita com as outras duas repetições, com exceção de uma região de 14 nucleotídeos idêntica que está presente apenas nas duas repetições maiores. (Figura 26 C)

6.5.3.2 Análise das proteínas de todos os transpovirons descritos por meio do *HHpred*

Após a caracterização genômica realizada para o Tvmarg, o próximo passo foi comparar os resultados obtidos com o que era predito para os outros seis transpovirons descritos. Dada a quantidade relativamente alta de proteínas sem função predita nos transpovirons, foram realizadas buscas por sequências homólogas no *HHpred* para cada uma das proteínas dos outros 6 transpovirons descritos.

Para os transpovirons do clado A, isto é, aqueles associados ao lentillevírus (Tvlent) e ao mamavírus (Tvmama), os resultados do *HHpred* foram os mesmos encontrados para o Tvmarg. Dessa forma, foram encontradas sequências homólogas para duas das quatro proteínas de Tvlent e Tvmama que estavam anotadas como proteínas hipotéticas, sendo elas a proteína com domínio transmembrana (tv_L2) e a possível transposase (tv_L4) (Tabela 12). Esses dados reforçam os domínios e funções que estão descritos para essas duas proteínas em trabalhos anteriores, apesar de não estarem anotados nos bancos de dados dessa forma.

No caso dos transpovirons do clado B, o que inclui as sequências relacionadas ao moumouvirus monve (Tvmonve) e moumouvirus australiensis (Tvmaust), para o Tvmonve foi possível encontrar sequências homólogas para duas das cinco proteínas anotadas como hipotéticas. A primeira delas é uma proteína envolvida com a regulação da transcrição (tv_R1)

e a segunda é uma proteína com motivo de zinco (tv_L8) (Tabela 13). Apesar de a função de regulação da transcrição já ter sido predita para a proteína tv_R1 em um trabalho anterior (JEUDY *et al.*, 2020), para a proteína tv_L8 não havia registro da detecção de sequências homólogas descrito anteriormente. Já para o Tvmaust, não foram encontradas sequências homólogas que pudessem alterar a anotação das proteínas. No que diz respeito aos transpovirons do clado C, ou seja, relacionados ao megavírus courdo7 (Tvcourdo) e megavírus vitis (Tvmvitis), não foram encontradas diferenças entre a anotação de suas proteínas e os resultados encontrados no *HHpred*. Os detalhes técnicos dessas análises podem ser encontrados no Anexo 3.

Tabela 13 – Resumo destacando as proteínas hipotéticas de transpovirons que apresentaram resultados diferentes da sua anotação, após busca por similaridade no *HHpred*.

	Transpoviron	Proteína	Nº de acesso (GenBank)	Resultado <i>HHpred</i>
Clado A de transpovirons	lentilevírus	tv_L2	AEY99248.1	Transmembrane protein: alpha-helical
	lentillevírus	tv_L4	AEY99250.1	Transposase (Uncharacterized protein Cgl2762)
	mamavírus	tv_L2	AEY99254.1	Transmembrane protein: alpha-helical
	mamavírus	tv_L4	AEY99256.1	Transposase (Uncharacterized protein Cgl2762)
Clado B de transpovirons	moumouvírus monve	tv_R1	AEY99259.1	Putative transcriptional regulator
	moumouvírus monve	tv_L8	AEY99266.1	zinc hook motif and AAA domain
	moumouvírus australiensis	-	-	sem diferenças
Clado C de transpovirons	megavírus courdo7	-	-	sem diferenças
	megavírus vitis	-	-	sem diferenças

6.5.3.3 Comparação entre o transpoviron do mimivírus argenteum e os outros transpovirons descritos

Quando se compara os genomas do transpoviron do mimivírus argenteum com os outros transpovirons descritos, é possível observar algumas diferenças no que diz respeito ao tamanho, ao conteúdo de G-C e ao número de ORFs. Quanto ao tamanho dos genomas, eles podem variar entre 6.594 até 7.584 pares de bases, sendo que a maior sequência pertence ao transpoviron do moumouvírus australiensis, enquanto a menor corresponde ao transpoviron do megavírus courdo7 (Tabela 14). Apesar disso, as sequências de transpoviron do moumouvírus monve e do megavírus courdo7 depositadas no banco de dados do NCBI são parciais. O conteúdo de G-C dos transpovirons é parecido e varia de 21,92% até 24,86%, representando uma média de 23,15%, similar a média de conteúdo de G-C de mimivírus (~26%). A diferença entre o conteúdo de G-C dos transpovirons do clado A para os do clado B é de, aproximadamente, 3%.

Já o número de ORFs varia entre 6 e 8, sendo que a maioria dos transpovirons apresenta 6 ORFs. Ao fazer uma comparação entre os três clados de transpovirons, o conteúdo de G-C e o número de ORFs é mais parecido nos transpovirons pertencentes ao mesmo clado (Tabela 14). Essa mesma diferença pode ser observada entre transpovirons do clado A e os dos clados B e C.

Outra análise realizada para fins comparativos foi o alinhamento da sequência de DNA do Tvmarg (*query*) contra as sequências de cada um dos outros transpovirons, por meio do *BLASTn*. Os resultados mostraram que o Tvmarg apresenta valores de identidade de nucleotídeos maiores quando alinhados com os transpovirons do lentillevírus (99,21%) e do mamavírus (99,11%), pertencentes ao clado A, sendo mais similar ao Tvlent (Tabela 14). Ao fazer o alinhamento da sequência do Tvmarg contra os transpovirons do clado B, foi possível observar valores de identidade de nucleotídeos de 72,52% e 73,88% para Tvmonve e Tvmaust, respectivamente. A diferença entre as sequências dos clados A e B está entre 26 e 27%, valor similar à diferença entre mimivírus de linhagem A e B. Em contrapartida, o alinhamento do Tvmarg com os transpovirons do clado C, utilizando essa metodologia, não indicou similaridade significativa entre eles (Tabela 14).

Tabela 14 – Resultados do alinhamento (*BLASTn*) da sequência de nucleotídeos do transpoviron do mimivírus *argentum* contra os outros transpovirons descritos e comparação de suas características genômicas.

Transpoviron - clado	Genoma (pb)	Conteúdo de G-C (%)	Número de ORFs	Identidade de nt com o Tvmarg (<i>BLASTn</i>)	Query cover	e-value
Mimivírus <i>argentum</i>	6.624	24,86	6	-	-	-
Lentillevírus - A	7.420	24,7	6	99,21%	100%	0.0
Mamavírus - A	7.438	24,59	6	99,11%	100%	0.0
Moumouvírus monve - B	6.717*	21,99	8	72,52%	24%	2,00E-142
Moumouvírus australiensis - B	7.584	22	8	73,88%	24%	2,00E-172
Megavírus courdo7 – C	6.594*	21,92	6	nenhuma similaridade significativa encontrada		
Megavírus vitis - C	7.417	22	7	nenhuma similaridade significativa encontrada		

Notas: pb: pares de bases; nt: nucleotídeos * Sequências incompletas

Em conjunto, os resultados dessa análise indicam que o Tvmarg é mais similar aos transpovirons do clado A. Além disso, também se observa que o Tvmarg apresenta maior similaridade com os transpovirons do clado B do que com aqueles do clado C. A similaridade entre Tvmarg e os transpovirons do clado A é reforçada na medida em que se analisa os valores de identidade de aminoácidos de cada proteína (Anexo 4). Essa comparação indica que todas

as seis proteínas do Tvmarg apresentam 100% de identidade de aminoácidos com as proteínas de transpovirons do clado A, sendo idêntica, inclusive, em relação ao tamanho da proteína ou do gene. Além de permitir a comparação das identidades de aminoácidos entre as proteínas do Tvmarg e as dos outros transpovirons descritos, o alinhamento entre proteínas de transpovirons no *BLASTp* também contribuiu para a realização de uma análise geral acerca da ocorrência dessas proteínas nos transpovirons, bem como da sintenia de seus genes, com base em possível homologia.

Dessa forma, foi possível concluir que os sete transpovirons descritos até o momento codificam, em conjunto, 10 proteínas diferentes. Dentre elas, seis são consideradas proteínas hipotéticas (enumeradas de 1 a 6), enquanto as outras 4 apresentam algum domínio ou função predita. Três dessas diferentes proteínas são conservadas e encontradas em todos os transpovirons, sendo elas a proteína hipotética 1, a proteína com domínio de helicase e a proteína com dedo de zinco. A alta identidade de aminoácidos (> 60%) entre Tvmarg e os outros transpovirons, observada para a proteína com domínio de helicase (T5), reforça a indicação dessa proteína como a mais conservada do grupo.

O Tvmarg apresenta os mesmos genes e a mesma sintenia, se comparado aos dois transpovirons do clado A (Figura 27). Além disso, ele compartilha com o Tvlent e com o Tvmama o gene para uma proteína que aparenta ser exclusiva do clado A, a proteína hipotética 2 (Figura 27). Os transpovirons do clado B são aqueles que possuem o maior número de genes (8) e uma maior diversidade de proteínas diferentes, em comparação aos outros. O Tvmaust apresenta 4 diferentes proteínas hipotéticas, sendo duas delas exclusivas (proteínas hipotéticas 3 e 4). Já Tvmonve apresenta duas proteínas hipotéticas (Figura 27). No caso dos transpovirons do clado C, o Tvmvitis codifica três proteínas hipotéticas, das quais uma é compartilhada apenas com o Tvmaust (proteína hipotética 6) e outra é compartilhada apenas com o Tvmonve (proteína hipotética 5) (Figura 27). Já o Tvcourdo7, apresentou apenas uma proteína hipotética, a proteína hipotética 1, presente em todos os outros transpovirons (Figura 27). Entretanto, o arquivo disponível para o Tvcourdo7 no banco de dados do NCBI era composto por apenas cinco sequências de aminoácidos, impossibilitando a análise completa desse transpoviron.

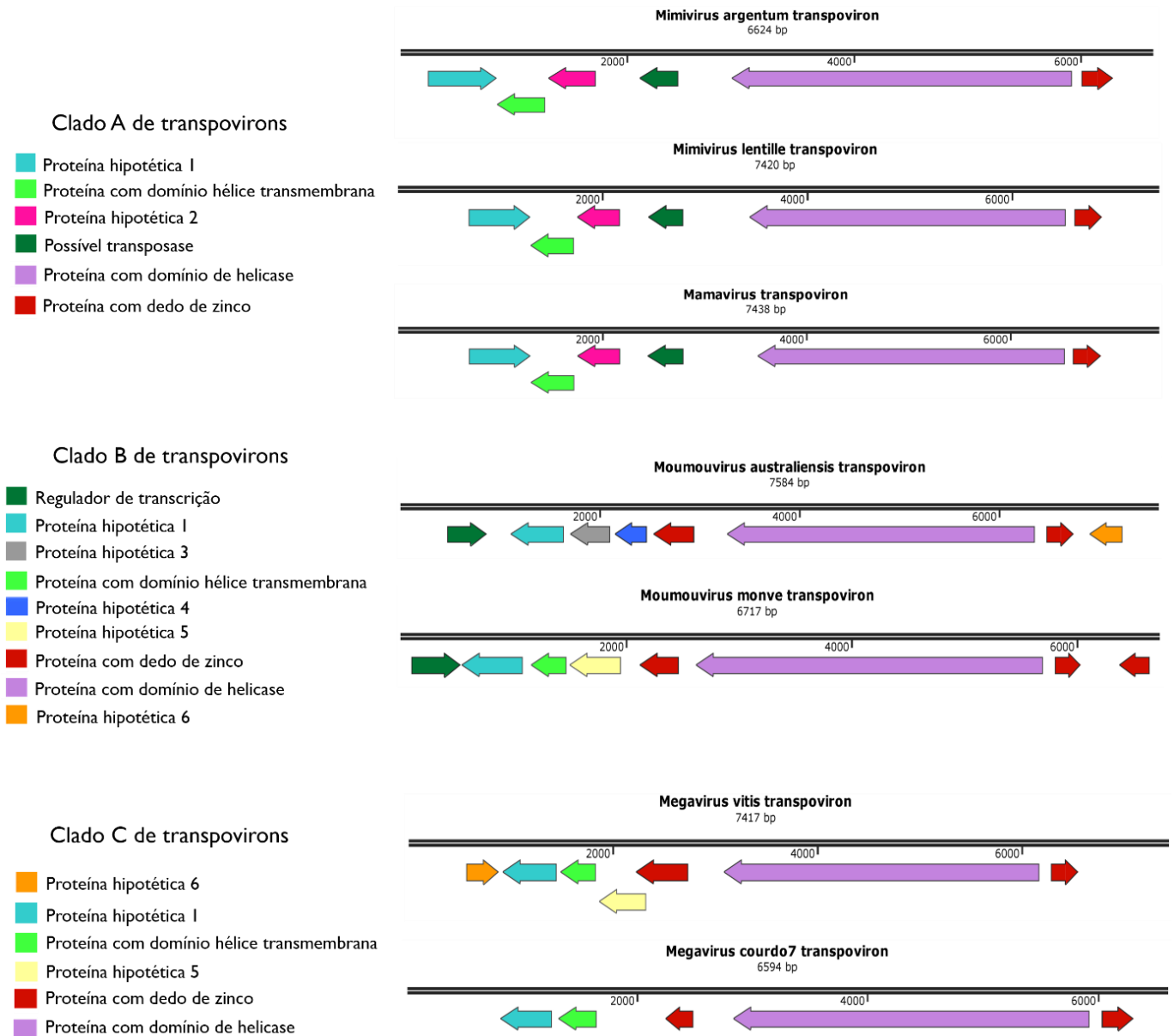


Figura 27 – Mapas individuais dos genomas dos transpovirons descritos, indicando a sintenia dos genes e as respectivas proteínas codificadas por cada um deles. As sequências homólogas estão representadas pelas mesmas cores. Obs.: O número de genes do transpoviron do megavírus courdo 7 é descrito nos trabalhos anteriores como 6, no entanto, no banco de dados do NCBI só foi possível encontrar 5 sequências. Obs.: As setas direcionadas para a direita indicam genes posicionados no sentido 5'-3' do DNA, enquanto as setas direcionadas para a esquerda indicam os genes posicionados no sentido 3'-5'.

Os clados A e B são os únicos a codificar uma proteína homóloga com função predita de transposase (clado A) ou reguladora de transcrição (clado B). Já a proteína com domínio hélice transmembrana (Tmhelix) é codificada por todos os transpovirons, com exceção do Tvmaust. A sintenia dos genes para as duas proteínas consideradas mais conservadas dentre os transpovirons, a proteína com domínio de helicase e a proteína com dedo de zinco, é parecida em todas as sequências (Figura 27). Com relação à proteína com dedo de zinco, nos clados B e

C, é possível observar a presença de parálogos, sendo que o Tvmone apresenta um parálogo a mais se comparado aos outros transpovirons (Figura 27).

Resumidamente, quando se leva em consideração o conteúdo de proteínas codificadas por pelo menos um transpoviron pertencente a cada clado, é possível concluir que o Clado A e o Clado B são os únicos a apresentar proteínas exclusivas e que eles compartilham uma proteína com função de transposase e/ou regulação da transcrição, que está ausente no clado C. Por sua vez, o clado C compartilha duas proteínas hipotéticas com o clado B. Por fim, também se observa a existência de quatro proteínas centrais, a proteína hipotética 1, a proteína com domínio de helicase, a proteína com dedo de zinco e a proteína com domínio transmembrana que são codificadas por todos os clados de transpovirons (Figura 28).

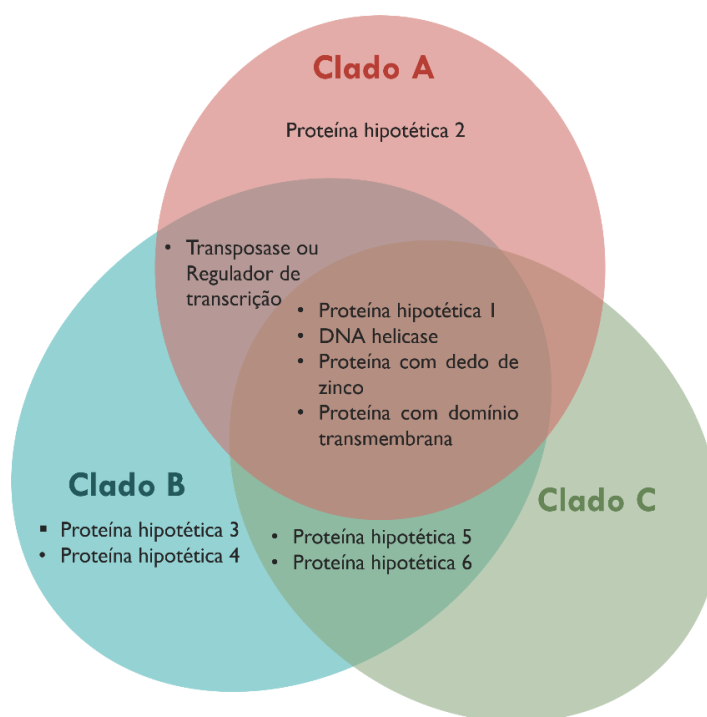


Figura 28 – Diagrama de Venn que resume a ocorrência das proteínas de transpovirons em cada um dos seus três clados (A, B e C).

6.5.3.4 Análises filogenéticas das proteínas do transpoviron do mimivírus argentinum

A comparação entre as proteínas do Tvmarg e o conjunto de proteínas de cada um dos outros transpovirons descritos também foi muito útil para análises filogenéticas. Dessa forma, foram construídas árvores filogenéticas para todas as proteínas do Tvmarg que estavam presentes em pelo menos dois clados. A análise filogenética da proteína hipotética 1 (T1), da proteína com domínio de helicase (T5) e da proteína com dedo de zinco (T6), que estão presentes em todos

os transpovirons descritos, corrobora com a divisão dos 3 clados (A, B e C), incluindo o Tvmarg no clado A de transpovirons e reforçando os dados apresentados anteriormente em outros trabalhos (Figura 29). Essa divisão de clados acompanha a separação das três linhagens de mimivírus, também chamada de A, B e C, uma vez que os transpovirons do clado A estão associados a mimivírus da linhagem A (Lentillevírus, mamavírus e mimivírus argentum), os transpovirons do clado B estão associados a mimivírus da linhagem B (moumouvírus monve e moumouvírus australiensis) e os transpovirons do clado C estão associados a mimivírus da linhagem C (megavírus courdo7 e megavírus vitis). No que diz respeito à árvore filogenética da proteína com dedo de zinco (T6), considerando que existem parálogos dessa proteína em transpovirons dos outros clados, para essa análise foram escolhidas apenas as sequências que eram os *best hits* da proteína T6 nas análises realizadas no *BLASTp*.

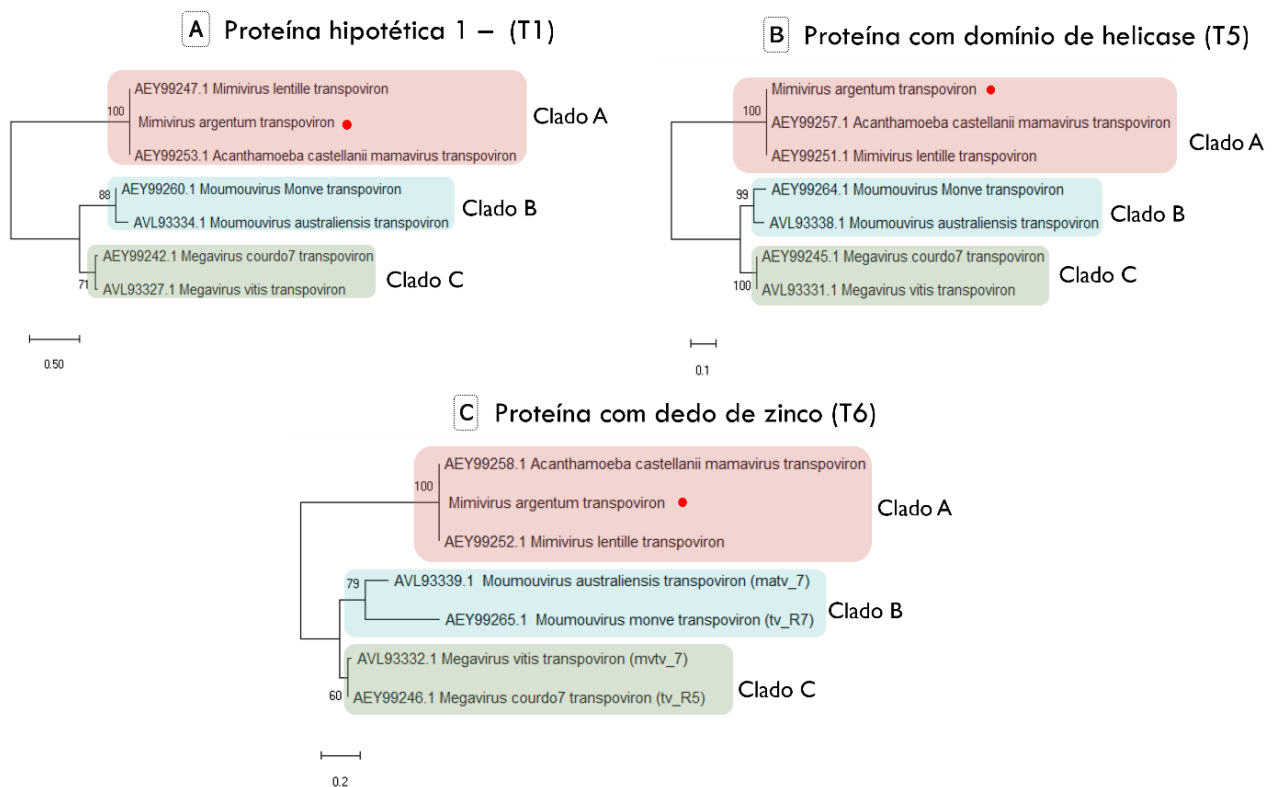


Figura 29 - Filogenia das proteínas T1, T5 e T6 do transpoviron do mimivírus argentum. (A-C) Árvores filogenéticas da proteína hipotética 1 (T1), proteína com domínio de helicase (T5) e da proteína com dedo de zinco (T6), respectivamente. Os alinhamentos foram realizados no MUSCLE, com parâmetros padrão. As árvores foram construídas no IQtree, utilizando o método de máxima verossimilhança, com suporte estatístico estimado com base em 1000 replicatas (*bootstrap*). Os melhores modelos, selecionados pelo IQtree (ModelFinder), para as árvores A, B e C foram JTTDCMut+F+G4, JTT+F+I e VT, respectivamente. As árvores construídas foram visualizadas no software MEGAX.

Além das proteínas T1, T5 e T6, foi possível realizar análises filogenéticas para outras duas proteínas do Tvmarg, a proteína com domínio transmembrana (T2) e a possível transposase (T4). A proteína com domínio transmembrana está presente em todos os transpovirons, com exceção do Tvmaust. A filogenia dessa proteína mostra que o clado A se mantém, entretanto, os outros dois cladros não são mantidos, sendo que o Tvmvitis (clado C) e o Tvmmonve (clado B) formam o mesmo ramo, compartilhando ancestral comum mais próximo. Já o Tvcourdo7 compõe um ramo divergente que surgiu a partir do mesmo ancestral que deu origem ao clado A (Figura 30 A).

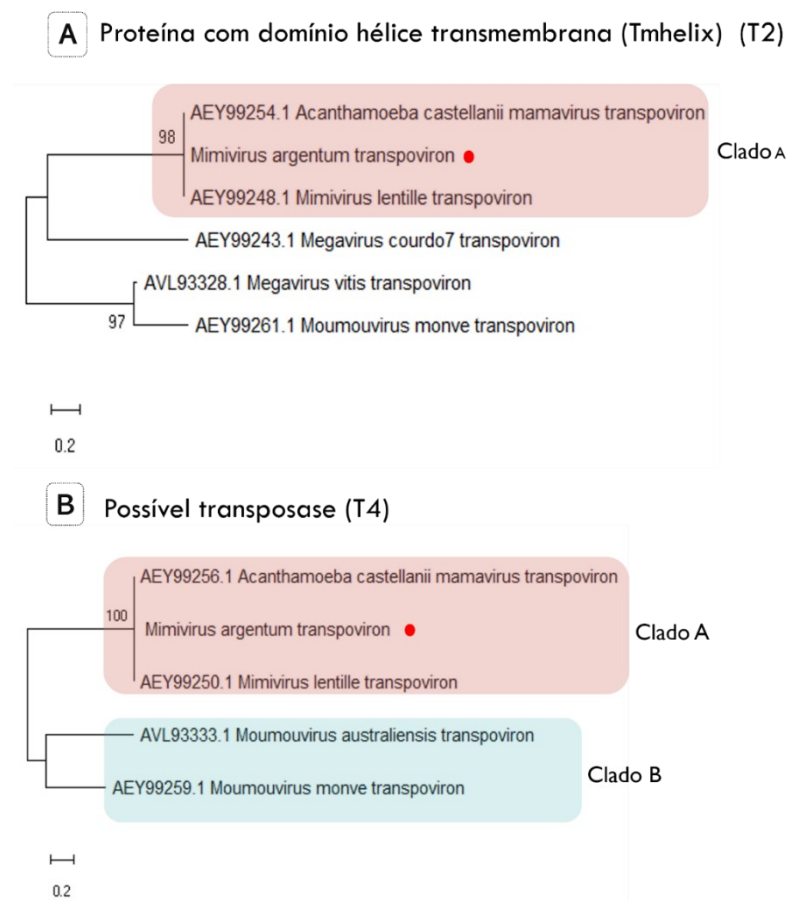


Figura 30 - Filogenia das proteínas T2 e T4 do transpoviron do mimivirus argentum. (A-B) Árvores filogenéticas da proteína com domínio hélice transmembrana (T2) e da possível transposase (T4), respectivamente. Os alinhamentos foram realizados no MUSCLE, com parâmetros padrão. As árvores foram construídas no IQtree, utilizando o método de máxima verossimilhança, com suporte estatístico estimado com base em 1000 replicatas (*bootstrap*). Os melhores modelos, selecionados pelo IQtree (ModelFinder), para as árvores A e B foram VT+F e Blosum62+F, respectivamente. As árvores construídas foram visualizadas no software MEGAX.

A filogenia da proteína T4, cujos homólogos estão presentes apenas nos transpovirons de clado A e B, mostra que ambos os clados se mantêm conservados (Figura 30 B). Para a proteína hipotética 2 (T3), não foram construídas árvores filogenéticas, uma vez que ela é exclusiva do clado A. Em conjunto, essas análises filogenéticas mostram que o clado A é bastante conservado, já que ele se manteve nas árvores de todas as proteínas do Tvmarg, o que não acontece com os outros dois clados. Isso reforça os resultados genômicos apresentados anteriormente, que indicam altos valores de identidade de aminoácidos entre as proteínas do Tvmarg e as dos transpovirons do clado A, além da similaridade na sintenia dos genes e da presença de uma proteína exclusiva do grupo (T3).

7 DISCUSSÃO

A primeira descrição dos mimivírus, em 2003, funcionou como um gatilho para a descoberta de inúmeros vírus gigantes, evidenciando a ampla diversidade e ubiquidade desses membros singulares da virosfera. Nesse contexto, o Brasil representa uma importante fonte de estudos envolvendo, não só a prospecção de novos vírus gigantes, como também a caracterização de seus isolados. O isolamento do mimivírus *argentum*, um mimivírus da linhagem A, a partir de uma lagoa urbana de Belo Horizonte - MG, vai de encontro com estudos que sugerem a ubiquidade e abundância desse tipo de vírus em amostras coletadas no Brasil (DORNAS *et al.*, 2015; ANDRADE *et al.*, 2018). Por outro lado, nós isolamos o terceiro virófago obtido a partir de amostras brasileiras, fato que, somado aos vários estudos com abordagem metagenômica que detectaram sequências de virófagos em diversos ambientes (ZHOU *et al.*, 2013; PAEZ-ESPINO *et al.*, 2019), sugere que as pesquisas visando o isolamento desses pequenos parasitas podem ser bem-sucedidas. Os estudos de prospecção de vírus gigantes e de virófagos se mostram importantes não somente devido ao isolamento viral em si, mas também porque facilitam a caracterização genômica, morfológica, estrutural e do ciclo de multiplicação desses vírus.

Neste trabalho, foi possível obter uma visão geral da morfologia das partículas isoladas, que, no caso dos mimivírus, auxiliou bastante na sua identificação. Utilizando a microscopia eletrônica de transmissão, observamos que o vírus gigante isolado apresentava características típicas de mimivírus, como o tamanho da partícula, a morfologia do capsídeo e a presença de fibrilas. De certa forma, isso também se aplica para o sputnik *argentum*. Embora as partículas de virófago sejam pequenas e não apresentem características tão marcantes como as de mimivírus, o fato de terem sido observadas em associação a fábricas virais, a vesículas ou a partículas do mimivírus, já funcionou como auxílio na identificação do sputnik *argentum*. Entretanto, vale ressaltar que as análises realizadas com base nas imagens de microscopia eletrônica, como as estimativas de tamanho de partícula, foram feitas com o simples objetivo de obter algumas características gerais das partículas isoladas. Assim, para uma caracterização mais aprofundada e detalhada seriam necessárias análises baseadas na utilização de métodos mais precisos, como a crio-microscopia eletrônica.

Quanto ao ciclo de multiplicação, as imagens de microscopia eletrônica de transmissão aqui apresentadas foram capazes de revelar determinadas características típicas do ciclo de mimivírus infectados por virófagos. Foi observada a utilização de recursos da fábrica viral do

vírus gigante para a formação de partículas de viróforo, como já descrito por trabalhos anteriores (DESNUES; RAOULT, 2010). Outra característica típica da coinfeção vírus gigante-viróforo, observada em nossas imagens, foi a formação de partículas defeituosas do vírus gigante, embora já tenha sido descrito que esse tipo de partícula também seja formado na ausência de viróforos (ANDRADE *et al.*, 2017). A observação de partículas de viróforos maduras, agrupadas ou dentro de vesículas, sugere que algumas etapas do ciclo de multiplicação dos viróforos precisam de estudos mais aprofundados. Apesar de a existência desses agrupamentos e de vesículas contendo vírions de viróforo já ter sido mostrada anteriormente (MOUGARI *et al.*, 2019a; JEUDY *et al.* 2019), mecanismos independentes de penetração ou liberação de partículas de viróforo que poderiam explicar essas observações ainda não foram descritos para os viróforos do gênero *Sputnik*. Dessa forma, estudos mais detalhados, como aqueles que consideram diferentes tempos pós-infecção ou o uso de inibidores específicos, podem ser valiosos para expandir o conhecimento sobre o assunto. Embora as imagens de microscopia eletrônica de transmissão tenham sido aliadas na identificação dos vírus isolados nesse trabalho, a maior parte da caracterização dos isolados aconteceu por meio dos estudos de genômica aqui apresentados.

A predição e anotação de genes realizadas para os três diferentes genomas estudados revelaram que eles são, de maneira geral, similares a outros genomas já descritos anteriormente. O genoma do mimivírus *argentum* se mostrou similar ao dos mimivírus da linhagem A. Os *best hits* do *BLASTp* para suas proteínas indicaram maior quantidade de *hits* com o mamavírus. Essa maior similaridade com as proteínas do mamavírus era esperada, já que a organização final dos *scaffolds* do mimivírus *argentum* foi realizada utilizando o genoma do mamavírus como referência. A relação com os mimivírus da linhagem A também pode ser reforçada pela filogenia, que refletiu a topologia dos clados de *Nucleocytoviricota* descrita anteriormente (IYER *et al.*, 2006; GUGLIELMINI *et al.*, 2019). Além de genes conservados, a anotação final das proteínas do mimivírus *argentum* permitiu identificar genes envolvidos em diferentes funções, como a tradução, o reparo de DNA e a participação em vias metabólicas específicas, que são típicos de mimivírus já descritos (RAOULT *et al.*, 2004; COLSON *et al.*, 2011). Outra categoria de proteínas encontrada em grandes proporções no genoma do mimivírus *argentum* é aquela composta por proteínas contendo repetições e domínios específicos. O grande número de proteínas desse tipo pode estar relacionado com o tamanho do genoma do mimivírus *argentum*, já que foi proposto que os vírus gigantes com genomas maiores, como é o caso dos mimivírus, tendem a ter uma maior quantidade dessas proteínas (SHUKLA *et al.*, 2018).

Sugere-se que essas proteínas contendo repetições e domínios tenham sido adquiridas por transferência gênica horizontal a partir de amebas e de bactérias, a qual foi facilitada por virófagos e transpovirons. Além disso, essas proteínas podem ter papel importante para o sucesso do ciclo de multiplicação dos mimivírus, uma vez que desempenham várias funções por meio de interações proteína-proteína (SHUKLA *et al.*, 2018).

Apesar da identificação de diferentes grupos de proteínas durante a anotação dos três genomas estudados, foram encontradas várias proteínas para as quais não se sabe nenhuma característica ou função, especialmente no mimivirus argenteus. A presença dessas proteínas, chamadas de proteínas hipotéticas, é reflexo da grande quantidade de ORFans presentes no genoma dos mimivírus e dos virófagos e transpovirons. Devido a essa presença marcante das ORFans nos genomas dos mimivírus, alguns trabalhos, utilizando metodologias diversas, já foram feitos com o intuito de revelar as funções das proteínas codificadas por algumas dessas ORFans (SOBHY *et al.*, 2015; SAINI *et al.*, 2007). O desenvolvimento desse tipo de pesquisa e constante adição de novas sequências nos bancos de dados podem tornar a anotação dessas proteínas hipotéticas desatualizada, especialmente considerando sequências depositadas há mais tempo.

Nesse contexto, novas buscas por sequências similares ou a predição de funções e estruturas podem ser estratégias promissoras. Dessa forma, utilizamos a ferramenta *HHpred*, com o objetivo de otimizar a anotação dos genomas estudados e predizer possíveis funções e/ou estruturas similares de sequências. Assim, foi possível comparar esses dados com a anotação que, usualmente, seria feita utilizando apenas o banco de dados de sequências de proteínas não redundantes do *BLASTp*. Conseguimos, assim, diminuir a quantidade de proteínas que seriam anotadas como hipotéticas para todos os três genomas aqui estudados, o que sugere que o uso de diferentes métodos e bancos de dados pode, de fato, contribuir para aprimorar a anotação do genoma de novos vírus isolados, bem como atualizar a anotação de vírus já descritos, especialmente os que apresentam um grande número de ORFans. Em contrapartida, a crescente descoberta de novos vírus gigantes acaba, por muitas vezes, expandindo ainda mais o conjunto de ORFans desses vírus. Isso pode ser exemplificado pela detecção de 12 ORFans exclusivas do mimivirus argenteus, para as quais não foi possível detectar sequências similares nem mesmo utilizando o *HHpred*. A identificação das funções das inúmeras proteínas hipotéticas dos vírus gigantes pode ajudar a direcionar estudos que atribuam essas proteínas a processos importantes do ciclo de multiplicação desses vírus, bem como suas interações com seus hospedeiros e parasitas. Portanto, as análises *in silico* mais acuradas, bem como diferentes experimentos

biológicos, também são essenciais para ajudar a desvendar funções e características escondidas por trás das ORFans dos vírus gigantes.

Em meio aos genes do mimivirus argenteum, também detectamos 20 ORFs que eram similares aos genes que compõem o genoma do virófago sputnik. As análises realizadas após o sequenciamento e a montagem, bem como as análises preliminares realizadas na ocasião do isolamento do vírus, sugeriram que o virófago sputnik argenteum seja capaz de se integrar no genoma do mimivírus argenteum. Os dados evidenciam, pela primeira vez, a possibilidade de integração do genoma completo de um provirófago no vírus gigante hospedeiro. A detecção de sequências endógenas contendo diversos genes de virófagos já havia sido descrita anteriormente, embora estivessem integradas nos genomas das células eucariotas hospedeiras (BLANC *et al.*, 2015; FISCHER; HACKL, 2016).

A análise da região onde o provirófago do sputnik argenteum está possivelmente inserido, se mostrou cercada por genes para proteínas contendo repetição de tripla hélice de colágeno, sendo que o mimivírus argenteum apresenta uma maior quantidade dessas proteínas, se comparado a outros mimivírus. Essas proteínas são codificadas tanto pelo vírus gigante, quanto pelo virófago, e as análises filogenéticas mostraram que algumas delas podem ter uma origem em comum. Em conjunto, esses achados vão de encontro com trabalho anterior que sugere a possibilidade de integração de provirófagos no genoma de mimivírus, a partir de regiões específicas de genes para proteínas com repetição de colágeno, podendo ser consideradas *hot spots* de integração de provirófago (DESNUES *et al.*, 2012). Caso regiões que contêm genes para essas proteínas sejam, de fato, *hot spots* de integração de provirófago, a presença de pequenos blocos de proteínas com repetição de colágeno localizados em outras partes do genoma, pode indicar a possibilidade de inserção do provirófago em diferentes sítios ao longo da mesma sequência hospedeira, corroborando com que fora proposto anteriormente.

Embora os mecanismos de integração e excisão dos provirófagos nos genomas hospedeiros ainda permaneçam desconhecidos, a presença do gene para uma integrase da família tirosina recombinase e de repetições diretas de 77 nt nas extremidades da sequência do sputnik argenteum, são fatores que podem direcionar estudos futuros com o objetivo de esclarecer esses mecanismos. As integrases são recombinases sítio-específicas que estão envolvidas na integração e excisão de EGMs em genomas hospedeiros e a presença de repetições diretas podem ser consideradas sítios para esse tipo de processo (WANG *et al.*, 2018; SEAH *et al.*, 2014). Contudo, é importante ressaltar que, apesar dos dados apresentados, é de

extrema importância que mais análises e experimentos biológicos diversos sejam feitos para suportar a existência tanto da integração do sputnik argentum em si, quanto desse possível sítio de inserção do provirófago no genoma do mimivirus argentum.

Em geral, o genoma do virófago sputnik argentum é mais similar ao dos virófagos sputnik 2 e rio negro. Entretanto, a análise das proteínas preditas para o sputnik argentum no *BLASTp*, revela que a maior parte de seus *best hits* são com o virófago sputnik 1 e que seu conteúdo de proteínas é muito parecido, com os genes apresentando uma sintenia também bastante similar. Isso pode ser reflexo dos altos valores de identidade de nucleotídeos compartilhados dentre os isolados de virófagos que compõem a mesma espécie dentro do gênero *Sputnikvirus*, indicando que eles são bastante similares entre si e diferem do sputnik argentum em uma porcentagem muito baixa. As análises filogenéticas realizadas reforçam essa ideia, pois agrupam o sputnik argentum com os outros isolados da espécie, mostrando uma baixa divergência entre as sequências, tanto nas árvores para as quatro proteínas mais conservadas da família, como na árvore para uma das proteínas contendo repetição de colágeno, presente no genoma do sputnik argentum. Além disso, a predição de estruturas realizada para as proteínas do sputnik argentum, por meio do *HHpred*, revelaram homologia com organismos, como bacteriófagos da classe *Tectiliviricetes* e adenovírus, que corroboram com o modelo de evolução proposto para os virófagos e vírus gigantes (KOONIN; KRUPOVIC; YUTIN, 2015).

Assim como o provirófago, também detectamos outro elemento do mobiloma dos mimivírus. Durante a anotação do genoma do transpoviron do mimivirus argentum, nós observamos que ele é mais similar geneticamente aos transpovirons do mamavírus e do lentillevírus, que compõem o clado A de transpovirons. As proteínas codificadas pelo Tvmarg também são codificadas pelo Tvmama e pelo Tvlent e apresentam valores altos de identidade de aminoácidos. Além disso, a sintenia dos genes desses três transpovirons é bastante similar, indicando a conservação desse clado, ideia que também é reforçada pela presença de uma proteína sem função predita que só é codificada pelas sequências de transpoviron do clado A. Apesar da similaridade do Tvmarg com os outros transpovirons do clado A, observamos que o Tvmama e o Tvlent são sequências maiores, com acréscimo de mais de 700 nucleotídeos. O Tvmarg também é menor do que os outros transpovirons dos clados B e C, com exceção do Tvmone e do Tvcourdo, cujas sequências depositadas no banco de dados são parciais.

Adicionalmente, observamos que as repetições terminais invertidas presentes nas extremidades do genoma do Tvmarg, típicas em EGMs, também são consideravelmente

menores do que as descritas para os outros transpovirons (DESNUES *et al.*, 2012; JEUDY *et al.*, 2019). Essas TIRs menores podem ser reflexo do menor tamanho do Tvmarg, já que parece ser improvável que a sua sequência esteja incompleta, considerando que a profundidade do seu sequenciamento foi relativamente alta (78x). A detecção dessas TIRs mais curtas no Tvmarg contrapõe a ideia proposta anteriormente de que as TIRs dos transpovirons são bem conservadas dentro dos clados (JEUDY *et al.*, 2019). Embora os dados aqui apresentados possam suportar essa proposição, um sequenciamento realizado com uma maior profundidade poderia corroborá-los. Apesar de ter sido mostrado que as TIRs de transpovirons possam ser formadas por repetições em tandem imperfeitas (DESNUES *et al.*, 2012), a existência de repetições diretas em uma das extremidades, como as que estão presentes na sequência do Tvmarg, não foram descritas anteriormente. Em geral, DNA transposons de eucariotos podem apresentar pequenas repetições diretas nas extremidades, chamadas de sítios de duplicação alvo, que são consideradas repetições derivadas da sequência onde se inseriram (KOJIMA, 2019). No caso do Tvmarg, as repetições diretas foram encontradas em apenas uma das extremidades e, sendo assim, a função ou participação dessas repetições no mecanismo de transposição e replicação dos transpovirons é ainda desconhecida. Deste modo, estudos mais aprofundados precisam ser desenvolvidos a fim de elucidar esses mecanismos nos transpovirons.

No que diz respeito à comparação realizada entre o conteúdo proteico de diferentes transpovirons, detectamos um conjunto de 10 proteínas diferentes que são codificadas pelos transpovirons conhecidos e três proteínas que são codificadas pelos *core genes* dos transpovirons, corroborando a descrição de trabalhos anteriores (JEUDY *et al.* 2019). Essas proteínas podem estar presentes em apenas um dos clados, em apenas dois clados ou em pelo menos um transpoviron de cada clado. Nossas análises mostraram que os transpovirons do clado B apresentam uma maior diversidade de proteínas e que compartilham, apenas com clado A, uma proteína com possível função de transposase ou regulação da transcrição. A ausência de ortólogos dessa proteína nos transpovirons do clado C é intrigante, reforçando a necessidade de estudos que visem entender os mecanismos de replicação e transposição desses EGMs. Apesar disso, é razoável hipotetizar que os transpovirons que não apresentam genes para essa enzima possam utilizar a transposase codificada pelos próprios mimivírus.

As análises comparativas entre o genoma do Tvmarg e o dos outros transpovirons mostraram, ainda, que as sequências que apresentam características genômicas mais similares, como conteúdo de G-C e número de ORFs, são pertencentes ao mesmo clado. Assim como os virófagos, os transpovirons apresentam valores médios de conteúdo de G-C similares aos de

mimivírus, apesar de ligeiramente menores. A diferença entre o conteúdo de G-C dos mimivírus da linhagem A e os das linhagens B e C também é similar ao que é observado nos clados de transpovirons correspondentes (~3%). Além disso, os valores de identidade de nucleotídeos obtidos quando se compara os transpovirons do clado A com os do clado B, são similares ao encontrado para os mimivírus da linhagem A e B. Além de agrupar o Tvmarg junto aos outros transpovirons do clado A, conforme o esperado, a filogenia das proteínas do Tvmarg revelou que a maioria delas é capaz de preservar a topologia dos clados de transpoviron, mesmo quando ausentes em alguma das sequências. Nas análises filogenéticas, o clado A se manteve em todas as proteínas analisadas, reforçando sua conservação. Em conjunto, as análises filogenéticas e comparativas apresentadas sugerem que a relação entre transpovirons e mimivírus, especialmente entre clado A de transpovirons e linhagem A de mimivírus, pode ser antiga e, provavelmente, anterior à divergência das linhagens de mimivírus. Um trabalho anterior, que testou experimentalmente essa afinidade dos transpovirons pelas linhagens específicas de mimivírus, também ajuda a reforçar essa hipótese (JEUDY *et al.*, 2019).

A contribuição dos elementos genéticos móveis para a transferência de genes entre diferentes organismos, torna o mobiloma dos mimivírus importantes personagens na história evolutiva dos vírus gigantes. Nossos resultados, em conjunto, sugerem que a descoberta e caracterização de novos vírus gigantes e de novos EGMs associados a eles são necessárias para expandir o conhecimento acerca da diversidade e distribuição desses elementos genéticos nos genomas dos vírus gigantes, bem como da interação vírus gigante-mobiloma, que pode estar intimamente relacionada com a evolução desses vírus. É notável que a relação entre os mimivírus e o seu mobiloma parece ser bastante complexa e fruto de uma longa história evolutiva, representando uma fonte importante de conhecimento que ainda precisa ser explorada.

8 CONCLUSÕES

- Descrevemos o isolamento de um mimivírus da linhagem A, a partir de amostras de uma lagoa urbana brasileira, reforçando a ubiquidade e abundância desses vírus no Brasil. Além disso, nós relatamos o isolamento do terceiro virófago obtido de amostras brasileiras, que é um virófago do gênero *Sputnikvirus*;
- Apresentamos uma visão geral acerca da morfologia das partículas isoladas e da interação entre elas, sugerindo que análises de microscopia eletrônica de transmissão podem ser aliadas na identificação de novos isolados;
- Foi possível obter a sequência do mimivirus argentum, após sequenciamento, e identificamos, por meio de análises *in silico*, a presença de dois elementos genéticos móveis associadas a ela: um transpoviron e um provirófago;
- O genoma do mimivirus argentum é similar ao dos outros mimivírus, codificando proteínas conservadas em *Nucleocyotoviricota*, bem como proteínas associadas a diferentes funções típicas de mimivírus. Além das proteínas hipotéticas, a maior parte das proteínas codificadas pelo mimivirus argentum são proteínas contendo repetições e/ou domínios, cujas funções são generalistas, como interações proteína-proteína;
- Os genes preditos para o sputnik argentum correspondem ao descrito para os outros virófagos de seu grupo, apresentando proteínas conservadas da família *Lavidaviridae* e outras proteínas com funções típicas desses organismos, incluindo aquelas que podem estar associadas com integração em genomas hospedeiros;
- O conteúdo proteico dos transpovirons é similar ao dos transpovirons pertencentes ao clado A, apresentando as três proteínas presentes em todos os transpovirons conhecidos, bem como uma proteína exclusiva dos transpovirons do clado A;
- A predição de domínios e/ou funções no *HHpred* conseguiu diminuir a quantidade de proteínas hipotéticas que seriam anotadas tanto para o mimivirus argentum, quanto para o transpoviron e para o sputnik argentum, uma vez que permitiu a predição de diferentes estruturas e funções para essas proteínas;

- Detectamos 20 ORFs no genoma do mimivirus argentum, com alta similaridade a proteínas de virófago, que estão cercadas por proteínas contendo repetição de colágeno, e indicam, pela primeira vez, a possibilidade de integração de genomas completos de provirófagos na sequência do vírus gigante hospedeiro. O mimivirus argentum apresenta um maior número de proteínas contendo repetição de tripla hélice de colágeno do que outros mimivirus, as quais podem representar *hot spots* de integração para o provirófago;
- A filogenia de uma das proteínas contendo repetição de colágeno do virófago sputnik argentum revelou uma relação evolutiva próxima com proteínas similares codificadas por mimivirus e homologia com outros membros de *Nucleocytoviricota*;
- A sequência do transpoviron do mimivirus argentum é menor do que as de seus similares, apresentando por consequência repetições terminais invertidas menores e contrapondo a ideia de conservação dessas repetições entre transpovirons do mesmo clado;
- A predição de estruturas e funções para as proteínas dos outros transpovirons descritos indicou diferenças entre o que estava anotado anteriormente para os transpovirons do lentillevírus e do mamavírus, bem como para o transpoviron do moumouvirus monve;
- As sequências de transpovirons descritos codificam 10 diferentes proteínas, estando três delas presentes em todas as sequências conhecidas. Os transpovirons do clado B apresentam uma maior diversidade de proteínas e compartilham proteínas exclusivas com cada um dos outros dois clados;
- As diferentes análises comparativas entre genomas dos transpovirons conhecidos, bem como as análises filogenéticas das proteínas do transpoviron do mimivirus argentum sugerem que o clado A de transpovirons seja o mais conservado dentre os outros três clados propostos;
- A relação de especificidade entre os três clados de transpovirons e as três linhagens do gênero *Mimivirus*, reforçada pelas análises filogenéticas, indicam que a relação entre transpovirons e mimivírus, especialmente entre clado A de transpovirons e linhagem A de mimivírus, pode ser antiga e que, provavelmente, se estabeleceu antes da divergência das linhagens de mimivírus.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABERGEL, C.; LEGENDRE, M.; CLAVERIE, J. The rapidly expanding universe of giant viruses: mimivirus, pandoravirus, pithovirus and mollivirus. **Fems Microbiology Reviews**, v.39, n.6, p.779-796, 20 set. 2015.
- ABRAHÃO *et al.* The analysis of translation-related gene set boosts debates around origin and evolution of mimiviruses. **Plos Genetics**, v. 13, n. 2, 16 fev. 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.1-12, 27 fev. 2018.
- ABRAHÃO, J. S. *et al.* Acanthamoeba polyphaga mimivirus and other giant viruses: an open field to outstanding discoveries. **Virology Journal**, v. 11, n. 120, 30 jun. 2014a
- ABRAHÃO, J. S. *et al.* Growing a giant: Evaluation of the virological parameters for mimivirus production. **Journal of Virological Methods**, v. 207, p.6-11, out. 2014b
- ANDRADE, A.C.S.P *et al.* Filling Knowledge Gaps for Mimivirus Entry, Uncoating, and Morphogenesis. **Journal of Virology**, v. 91, n. 22, 6 set. 2017.
- ANDRADE, A.C.S.P. *et al.* Ubiquitous giants: a plethora of giant viruses found in Brazil and Antarctica. **Virology Journal**, v. 15, n. 1, 24 jan. 2018
- ANDRADE, A.C.S.P.A *et al.* New Isolates of Pandoraviruses: Contribution to the Study of Replication Cycle Steps. **Journal of Virology**, v.93, n.5, 19 fev. 2019.
- ANDREANI, J. *et al.* Cedratvirus, a Double-Cork Structured Giant Virus, is a Distant Relative of Pithoviruses. **Viruses**, v. 8, n. 11, p.300-311, 3 nov. 2016
- ANDREANI, J. *et al.* Orpheovirus IHUMI-LCC2: A New Virus among the Giant Viruses. **Frontiers in Microbiology**, v. 8, 22 jan. 2018.
- ANDREANI, J. *et al.* Pacmanvirus, a New Giant Icosahedral Virus at the Crossroads between Asfarviridae and Faustoviruses. **Journal of Virology**, v. 91, n. 14, p.1-11, 26 abr. 2017.
- ARANTES, T.S. *et al.* The Large Marseillevirus Explores Different Entry Pathways by Forming Giant Infectious Vesicles. **Journal of Virology**, v. 90, n. 11, p. 5246-5255, 16 mar. 2016.
- 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. 17486-91. 18 out. 2011
- BAJRAI, L. *et al.* Kaumobavirus, a New Virus That Clusters with Faustoviruses and Asfarviridae. **Viruses**, v. 8, n. 11, p.278-288, 28 out. 2016.
- BAJRAI, L. H. *et al.* Isolation of Yasminevirus, the First Member of Klosneuvirinae Isolated in Coculture with Vermamoeba vermiformis, Demonstrates an Extended Arsenal of Translational Apparatus Components. **Journal of Virology**, v. 94, n. 1, 9 out. 2019.
- BAJRAI, L.H. *et al.* Saudi Moumouvirus, the First Group B Mimivirus Isolated from Asia. **Frontiers in Microbiology**, v. 07, 20 dez. 2016.

- BANKEVICH, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455-477, maio 2012.
- BARIK, S. A Family of Novel Cyclophilins, Conserved in the Mimivirus Genus of the Giant DNA Viruses. **Computational and Structural Biotechnology Journal**, v. 16, p. 231-236, 2018.
- BEKLIZ, M. *et al.* A New Zamilon-like Virophage Partial Genome Assembled from a Bioreactor Metagenome. **Frontiers in Microbiology**, v. 6, 27 nov. 2015.
- BEKLIZ, M.; COLSON, P.; LA SCOLA, B. The Expanding Family of Virophages. **Viruses**, v. 8, n. 11, 23 nov. 2016
- BELLAS, C. M.; ANESIO, A. M.; BARKER, G. Analysis of virus genomes from glacial environments reveals novel virus groups with unusual host interactions. **Frontiers in Microbiology**, v. 6, 3 jul. 2015.
- BESEMER, J. *et al.* GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. **Nucleic Acids Research**, v. 33, p. 451-454, 1 jul. 2005
- BLANC, G. al. Provirophages in the Bigelowiella genome bear testimony to past encounters with giant viruses. **Proceedings of The National Academy of Sciences**, v. 112, n. 38, p. 5318-5326, 24 ago. 2015.
- BLUM, M. *et al.* The InterPro protein families and domains database: 20 years on. **Nucleic Acids Research**, v. 49, n. 1, p. 344-354, 6 nov. 2020.
- 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, 10 nov. 2015.
- BORATTO, P.V. M. *et al.* Yaravirus: a novel 80-nm virus infecting *Acanthamoeba castellanii*. **Proceedings of The National Academy of Sciences**, v. 117, n. 28, p. 16579-16586, 29 jun. 2020.
- BORGES, I. A. *et al.* Rio Negro virophage: Sequencing of the near complete genome and transmission electron microscopy of viral factories and particles. **Brazilian Journal of Microbiology**, v. 49, p.260-261, nov. 2018.
- BORGES, I. *et al.*, Trapping the Enemy: *Vermamoeba vermiformis* Circumvents Faustovirus Mariensis Dissemination by Enclosing Viral Progeny inside Cysts. **Journal of Virology**, v.93, n.14, 28 jun. 2019
- BORN, D. *et al.* Capsid protein structure, self-assembly, and processing reveal morphogenesis of the marine virophage mavirus. **Proceedings of The National Academy of Sciences**, v. 115, n. 28, p. 7332-7337, 25 jun. 2018.
- BOSI, E. *et al.* MeDuSa: a multi-draft based scaffolder. **Bioinformatics**, v. 31, n. 15, p. 2443-2451, 25 mar. 2015.

- BOUGHALMI, M. *et al.* High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. **Environmental Microbiology**, v. 15, n. 7, p. 2000-2007, 9 jan. 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, 22 dez. 2009.
- BOYER, M. *et al.* Mimivirus shows dramatic genome reduction after intraamoebal culture. **Proceedings of The National Academy of Sciences**, v. 108, n. 25, p.10296-10301, 6 jun. 2011.
- BOYER, M. *et al.* Phylogenetic and Phyletic Studies of Informational Genes in Genomes Highlight Existence of a 4th Domain of Life Including Giant Viruses. **Plos One**, v. 5, n. 12, 2 dez. 2010.
- 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.95-105, 2014.
- CHAN, P. P. *et al.* tRNAscan-SE: searching for tRNA genes in genomic sequences. **Methods in Molecular Biology**, v. 1962, p. 1-14, 2019.
- CHATTERJEE, A. *et al.* Genomic and metagenomic signatures of giant viruses are ubiquitous in water samples from sewage, inland lake, wastewater treatment plant, and municipal water supply in Mumbai, India. **Scientific Reports**, v. 9, n. 1, 6 mar. 2019
- CHRISTO-FOROUX, E. *et al.* Characterization of Mollivirus kamchatka, the First Modern Representative of the Proposed Molliviridae Family of Giant Viruses. **Journal of Virology**, v. 94, n. 8, 29 jan. 2020.
- CLAVERIE J.M., ABERGEL C., OGATA H. Mimivirus. In: VAN ETTEN, J. L. *et al.* (ed.). **Lesser Known Large dsDNA Viruses**. Berlin, Heidelberg: Springer, 2009. p. 89-121. (Current Topics in Microbiology and Immunology, v.328).
- CLAVERIE, J- M.; ABERGEL, C. Mimiviridae: An Expanding Family of Highly Diverse Large dsDNA Viruses Infecting a Wide Phylogenetic Range of Aquatic Eukaryotes. **Viruses**, v. 10, n. 9, p.506-519, 18 set. 2018.
- CLAVERIE, J.M.; ABERGEL, C. CRISPR-Cas-like system in giant viruses: why MIMIVIRE is not likely to be an adaptive immune system. **Virologica Sinica**, v. 31, n. 3, p.193-196, jun. 2016.
- COLSON, P. *et al.* "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. **Archives of Virology**, v. 158, n. 12, p.2517-2521, 29 jun. 2013.
- COLSON, P. *et al.* Giant mimiviruses escape many canonical criteria of the virus definition. **Clinical Microbiology and Infection**, v. 25, n. 2, p. 147-154, 2018
- COLSON, P. *et al.* Mimivirus: leading the way in the discovery of giant viruses of amoebae. **Nature Reviews Microbiology**, v. 15, n. 4, p. 243-254, abr. 2017.
- COLSON, P. *et al.* Reclassification of Giant Viruses Composing a Fourth Domain of Life in the New Order Megavirales. **Intervirology**, v. 55, n. 5, p.321-332, 2012.

COLSON, P. *et al.* Viruses with More Than 1,000 Genes: mamavirus, a new *Acanthamoeba polyphaga mimivirus* strain, and reannotation of mimivirus genes. **Genome Biology and Evolution**, v. 3, p. 737-742, 1 jan. 2011.

DESNUES, C. *et al.* Provirophages and transpovirons as the diverse mobilome of giant viruses. **Proceedings of The National Academy of Sciences**, v. 109, n. 44, p.18078-18083, 15 out. 2012.

DESNUES, C.; RAOULT, D. Inside the Lifestyle of the Virophage. **Intervirology**, v. 53, n. 5, p. 293-303, 2010.

DESNUES, C.; RAOULT, D. Virophages question the existence of satellites. **Nature reviews Microbiology**. 2012

DESNUES, C.; BOYER, M.; RAOULT, D. Sputnik, a Virophage Infecting the Viral Domain of Life. **Bacteriophages, Part A**, p. 63-89, 2012.

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.1086, 06 out. 2015.

DOU, C. *et al.* Structural and Mechanistic Analyses Reveal a Unique Cas4-like Protein in the Mimivirus Virophage Resistance Element System. **Iscience**, v. 3, p.1-10, maio 2018.

DUPONCHEL, S.; FISCHER, M.G. Viva lavidaviruses! Five features of virophages that parasitize giant DNA viruses. **Plos Pathogens**, v. 15, n. 3, 21 mar. 2019.

EDGAR, R. C. *et al.* MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792-1797, 8 mar. 2004.

FESCHOTTE, C.; PRITHAM, E. J. DNA Transposons and the Evolution of Eukaryotic Genomes. **Annual Review of Genetics**, v. 41, n. 1, p.331-368, dez. 2007.

FILÉE, J. Giant viruses and their mobile genetic elements: the molecular symbiosis hypothesis. **Current Opinion in Virology**, v. 33, p. 81-88, dez. 2018

FILÉE, J.; CHANDLER, M. Gene Exchange and the Origin of Giant Viruses. **Intervirology**, v. 53, n. 5, p. 354-361, 2010.

FISCHER, M. G. Sputnik and Mavirus: more than just satellite viruses. **Nature Reviews Microbiology**, v. 10, n. 1, p.78-78, 5 dez. 2011.

FISCHER, M. G.; SUTTLE, C. A. A Virophage at the Origin of Large DNA Transposons. **Science**, v. 332, n. 6026, p.231-234, 3 mar. 2011.

FISCHER, M.; HACKL, T. Host genome integration and giant virus-induced reactivation of the virophage mavirus. **Nature**, v. 540, n. 7632, p. 288-291, 7 dez. 2016.

FISCHER, M.G. The Virophage Family *Lavidaviridae*. **Current Issues in Molecular Biology**, v.40, p. 1-24, 2020

FROST, L. S. *et al.* Mobile genetic elements: the agents of open source evolution. **Nature Reviews Microbiology**, v. 3, n. 9, p.722-732, set. 2005.

GAIA, M. *et al.* Broad Spectrum of Mimiviridae Virophage Allows Its Isolation Using a Mimivirus Reporter. **Plos One**, v. 8, n. 4, 15 abr. 2013a.

GAIA, M. *et al.* The Virophage Concept. In: **eLS**. John Wiley & Sons, Ltd: Chichester. 15 maio 2013b.

GAIA, M. *et al.* Zamilon, a Novel Virophage with Mimiviridae Host Specificity. **Plos One**, v. 9, n. 4, 18 abr. 2014.

GONG, C. *et al.* Novel Virophages Discovered in a Freshwater Lake in China. **Frontiers in Microbiology**, v. 7, 22 jan. 2016.

GUGLIELMINI, J. *et al.* Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. **Proceedings of The National Academy of Sciences**, v. 116, n. 39, p. 19585-19592, 10 set. 2019.

GULINO, K. *et al.* Initial Mapping of the New York City Wastewater Virome. **mSystems**, v. 5, n. 3, 16 jun. 2020.

HAKIM, M. *et al.* Exploring ORFan Domains in Giant Viruses: structure of mimivirus sulfhydryl oxidase r596. **Plos One**, v. 7, n. 11, 28 nov. 2012.

ICTV ONLINE. **Taxonomy**. 2020. Disponível em: <<https://talk.ictvonline.org/taxonomy/>>. Acesso em: 31 dez. 2020

IYER, L. M. *et al.* Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. **Virus Research**, v. 117, n. 1, p.156-184, abr. 2006.

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

JEUDY, S. *et al.* Exploration of the propagation of transpovirons within Mimiviridae reveals a unique example of commensalism in the viral world. **The ISME Journal**, v. 14, n. 3, p. 727-739, 10 dez. 2019.

KALYAANAMOORTHY, S. *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. **Nature Methods**, v. 14, n. 6, p. 587-589, 8 maio 2017.

KAPITONOV, V.V.; JURKA, J. Rolling-circle transposons in eukaryotes. **Proceedings of The National Academy of Sciences**, v. 98, n. 15, p.8714-8719, 10 jul. 2001.

KHALIL, J. Y.B.; ANDREANI, J.LA SCOLA, B. Updating strategies for isolating and discovering giant viruses. **Current Opinion in Microbiology**, v.31, p. 80-87, Jun. 2016.

KIDWELL, M.G.; LISCH, D. R. Perspective: transposable elements, parasitic DNA, and genome evolution. **International Journal of Organic Evolution**, v. 55, n. 1, p.1-24, jan. 2001.

- KLOSE, T. *et al.* The Three-Dimensional Structure of Mimivirus. **Intervirology**, v. 53, n. 5, p.268-273, 2010.
- KOJIMA, Kenji K. Structural and sequence diversity of eukaryotic transposable elements. **Genes & Genetic Systems**, v. 94, n. 6, p. 233-252, 1 dez. 2019.
- KOONIN, E. V. Viruses and mobile elements as drivers of evolutionary transitions. **Philosophical Transactions of The Royal Society B: Biological Sciences**, v. 371, n. 1701, 18 jul. 2016.
- KOONIN, E. V.; YUTIN, N. Evolution of the Large Nucleocytoplasmic DNA Viruses of Eukaryotes and Convergent Origins of Viral Gigantism. **Advances in Virus Research**, p.167-202, 2019.
- KOONIN, E. V; KRUPOVIC, M. Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity. *Current Opinion in Virology*, v. 25, p.7-15, ago. 2017.
- KOONIN, E.; KRUPOVIC, M.; YUTIN, N. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. **Annals of The New York Academy of Sciences**, v. 1341, n. 1, p. 10-24, 27 fev. 2015.
- KOONIN, E.V. *et al.* Global Organization and Proposed Megataxonomy of the Virus World. **Microbiology and Molecular Biology Reviews**, v. 84, n. 2, 4 mar. 2020
- KOONIN, E.V. Virology: gulliver among the lilliputians. **Current Biology**, v. 15, n. 5, p. 167-169, mar. 2005.
- KRUPOVIC, M.; CVIRKAITE-KRUPOVIC, V. Virophages or satellite viruses? **Nature Reviews Microbiology**, v. 9, n. 11, p. 762-763, 11 out. 2011.
- KRUPOVIC, M.; KOONIN, E. V. Self-synthesizing transposons: unexpected key players in the evolution of viruses and defense systems. **Current Opinion in Microbiology**, v. 31, p. 25-33, jun. 2016.
- KRUPOVIC, M; KUHN, J.H.; FISCHER, M.G. A classification system for virophages and satellite viruses. **Archives of Virology**, v. 161, n. 1, p. 233-247, 7 out. 2015.
- KUMAR, S. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547-1549, 2 maio 2018.
- 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, 7 ago. 2013.
- LA SCOLA, B. *et al.* A Giant Virus in amoebae. **Science**. 299: 2033. 2003
- LA SCOLA, B. *et al.* The virophage as a unique parasite of the giant mimivirus. **Nature**, v. 455, n. 7209, p. 100-104, 6 ago. 2008.

- LAMB, D. C. *et al.* The First Virally Encoded Cytochrome P450. **Journal of Virology**, v. 83, n. 16, p. 8266-8269, 10 jun. 2009.
- LASLETT, D. *et al.* ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. **Nucleic Acids Research**, v. 32, n. 1, p. 11-16, 2 jan. 2004.
- LEGENDRE, M. *et al.* Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. **Virology Journal**, v. 8, n. 1, 4 mar. 2011.
- 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.5327-5335, 8 set. 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, 3 mar. 2014.
- LETUNIC, I. *et al.* Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. **Bioinformatics**, v. 23, n. 1, p. 127-128, 18 out. 2006.
- LEVASSEUR, A. Comparison of a Modern and Fossil Pithovirus Reveals Its Genetic Conservation and Evolution. **Genome Biology and Evolution**, v. 8, n. 8, p. 2333-2339, 7 jul. 2016.
- LEVASSEUR, A. *et al.* MIMIVIRE is a defense system in mimivirus that confers resistance to virophage. **Nature**, v. 531, n. 7593, p.249-252, 29 fev. 2016.
- LEVIN, H.L.; MORAN, J. V. Dynamic interactions between transposable elements and their hosts. **Nature Reviews Genetics**, v. 12, n. 9, p.615-627, 18 ago. 2011.
- LUTHER, K.B. *et al.* Mimivirus Collagen Is Modified by Bifunctional Lysyl Hydroxylase and Glycosyltransferase Enzyme. **Journal of Biological Chemistry**, v. 286, n. 51, p. 43701-43709, 1 nov. 2011
- LWOFF, A. The Concept of Virus. **Journal of General Microbiology**. v.17, n.1, p. 239-253. 23 ago. 1957g
- MCCLINTOCK, B. The origin and behavior of mutable loci in maize. **Proceedings of the National Academy of Sciences**. v. 36, p. 344–355, 1950.
- MICHEL. R. *et al.* Experimental co-infection of *Saccamoeba lacustris* with Mimivirus-like Giant virus and a small Satellite virus. **Endocytobiosis and Cell Research**, v. 29, p. 1-6, Maio 2018.
- MILLER, W.J.; CAPY, P. Mobile Genetic Elements as Natural Tools for Genome Evolution. **Methods in Molecular Biology**, v. 260, p. 001-020, 2004.
- MOUGARI, S. *et al.* Guarani Virophage, a New Sputnik-Like Isolate From a Brazilian Lake. **Frontiers in Microbiology**, v. 10, 3 maio 2019a.
- MOUGARI, S. *et al.* Role of the R349 Gene and Its Repeats in the MIMIVIRE Defense System. **Frontiers in Microbiology**, v. 10, 22 maio 2019b.

MOUGARI, S. *et al.* Virophages of Giant Viruses: an update at eleven. **Viruses**, v. 11, n. 8,8 ago. 2019c.

MUTSAFI, Y. *et al.* Membrane Assembly during the Infection Cycle of the Giant Mimivirus. **Plos Pathogens**, v. 9, n. 5, 30 maio 2013.

MUTSAFI, Y. *et al.* Vaccinia-like cytoplasmic replication of the giant Mimivirus. **Proceedings of The National Academy of Sciences**, v. 107, n. 13, p.5978-5982, 15 mar. 2010.

MYLLYHARJU, J *et al.* Collagens, modifying enzymes and their mutations in humans, flies and worms. **Trends in Genetics**, v. 20, n. 1, p. 33-43, jan. 2004.

NELSON, D.L; COX, M.M. Estrutura tridimensional das proteínas. In: NELSON, D.L; COX, M.M. **Princípios de Bioquímica de Lehninger**. 7. ed. Porto Alegre: Artmed, 2019. Cap. 4. p. 127-130.

NGUYEN, L.T. *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. **Molecular Biology and Evolution**, v. 32, n. 1, p. 268-274, 3 nov. 2014.

OGATA, H.; RAOULT, D.; CLAVERIE, J-M. A new example of viral intein in Mimivirus. **Virology Journal**, v. 2, n. 1, p. 8, 2005.

OH, S.; YOO, D.; LIU, W. Metagenomics Reveals a Novel Virophage Population in a Tibetan Mountain Lake. **Microbes and Environments**, v. 31, n. 2, p.173-177, 2016.

PAEZ-ESPINO, D. *et al.* Diversity, evolution, and classification of virophages uncovered through global metagenomics. **Microbiome**, v. 7, n. 1, dez. 2019.

PAGNIER, I. *et al.* A Decade of Improvements in Mimiviridae and Marseilleviridae Isolation from Amoeba. **Intervirology**, v. 56, n. 6, p.354-363, 2013.

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, 18 jul. 2013.

PIACENTE, F.*et al.* The Autonomous Glycosylation of Large DNA Viruses. **International Journal of Molecular Sciences**, v. 16, n. 12, p. 29315-29328, 9 dez. 2015.

PRITHAM, E.J.; PUTLIWALA, T.; FESCHOTTE, C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. **Gene**, v. 390, n. 1-2, p. 3-17, abr. 2007.

PRJIBELSKI, A. *et al.* Using SPAdes De Novo Assembler. **Current Protocols In Bioinformatics**, v. 70, n. 1, jun. 2020.

RAOULT, D. *et al* The 1.2-Megabase Genome Sequence of Mimivirus. **Science**, v. 306, n. 5700, p.1344-1350, 19 nov. 2004.

RAOULT, D.; BOYER, M. Amoebae as Genitors and Reservoirs of Giant Viruses. **Intervirology**, v. 53, n. 5, p.321-329, 2010.

RAOULT, D.; BOYER, M. Amoebae as Genitors and Reservoirs of Giant Viruses. **Intervirology**, v. 53, n. 5, p. 321-329, 2010.

RAOULT, D.; FORTERRE, P. Redefining viruses: lessons from Mimivirus. **Nature Reviews Microbiology**, v. 6, n. 4, p.315-319, 3 mar. 2008.

RAOULT, D.; LA SCOLA, B.; BIRTLES, R. The Discovery and Characterization of Mimivirus, the Largest Known Virus and Putative Pneumonia Agent. **Clinical Infectious Diseases**. v. 45, n. 1, p. 95-102, 1 jul. 2007.

RAVINDRAN, S. Barbara McClintock and the discovery of jumping genes. **Proceedings of The National Academy of Sciences**, v. 109, n. 50, p. 20198-20199, 11 dez. 2012.

REED, L.J.; MUENCH, H. A Simple method of estimating fifty per cent endpoints. **The American Journal of Hygiene**. v. 27. n. 3, p. 493-497, 1938.

RENESTO, P. *et al.* Mimivirus Giant Particles Incorporate a Large Fraction of Anonymous and Unique Gene Products. **Journal of Virology**, v. 80, n. 23, p. 11678-11685, 13 set. 2006.

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, 15 abr. 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, 16 set. 2015.

RODRIGUES, R.A.L *et al.* Morphologic and Genomic Analyses of New Isolates Reveal a Second Lineage of Cedratviruses. **Journal of Virology**, v. 92, n. 13, p. 1-13, 25 abr. 2018a.

RODRIGUES, R. A. L. *et al.* “Tupanvirus”, a new genus in the family Mimiviridae. **Archives of Virology**, v. 164, n. 1, p.325-331, 5 out. 2018b

RODRIGUES, R.A.L.; SILVA, L.C.F.; ABRAHÃO, J.S. Translating the language of giants: translation-related genes as a major contribution of giant viruses to the virosphere. **Archives of Virology**, v. 165, n. 6, p. 1267-1278, 24 abr. 2020.

ROLLAND, C. *et al.* Discovery and Further Studies on Giant Viruses at the IHU Mediterranean Infection That Modified the Perception of the Virosphere. **Viruses**, v. 11, n. 4, 30 mar. 2019.

ROUX, S. *et al.* Ecogenomics of virophages and their giant virus hosts assessed through time series metagenomics. **Nature Communications**, v. 8, n. 1, 11 out. 2017

SAINI, H.*et al.* Structural and functional insights into Mimivirus ORFans. **Bmc Genomics**, v. 8, n. 1, p. 115-118, 2007.

SANTINI, S. *et al.* Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. **Proceedings of The National Academy of Sciences**, v. 110, n. 26, p.10800-10805, 10 jun. 2013

SCHNEIDER, C. A.; RASBAND, W. S.; ELICEIRI, K. W. NIH Image to ImageJ: 25 years of image analysis. **Nature methods**, v. 9, n. 7, p.671-675, 2012.

SCHRAD, J.R. *et al.* Structural and Proteomic Characterization of the Initiation of Giant Virus Infection. **Cell**, v. 181, p. 1046–1061. 28 maio 2020.

SCHRADER, L.; SCHMITZ, J. The impact of transposable elements in adaptive evolution. **Molecular Ecology**, v. 28, n. 6, p.1537-1549, 4 ago. 2018.

SCHULZ, F. *et al.* Giant viruses with an expanded complement of translation system components. **Science**, v. 356, p. 82-85, 7 abr. 2017

SHAH, N. *et al.* Exposure to Mimivirus Collagen Promotes Arthritis. **Journal of Virology**, v. 88, n. 2, p. 838-845, jan. 2014

SHUKLA, A. *et al.* The number of genes encoding repeat domain-containing proteins positively correlates with genome size in amoebal giant viruses. **Virus Evolution**, v. 4, n. 1, 1 jan. 2018.

SIEFERT, J. L. Defining the Mobilome. **Methods in Molecular Biology**, p.13-27, Fev. 2009.

SILVA, L.K.S. *et al.* Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. **Scientific Reports**, v. 8, n. 1, p. 1-12, 5 mar. 2018.

SILVA, L.K.S. *et al.* Isolation and genomic characterization of a new mimivirus of lineage B from a Brazilian river. **Archives of Virology**, v. 165, n. 4, p. 853-863, 12 fev. 2020.

SOBHY, H. *et al.* Identification of giant Mimivirus protein functions using RNA interference. **Frontiers in Microbiology**, v. 6, 28 abr. 2015.

SODING, J. *et al.* The HHpred interactive server for protein homology detection and structure prediction. **Nucleic Acids Research**, v. 33, p. 244-248, 1 jul. 2005

STOUGH, J. M. A. *et al.* Genome and Environmental Activity of a Chrysochromulina parva Virus and Its Virophages. **Frontiers in Microbiology**, v. 10, 5 abr. 2019.

SUBRAMANIAM, K. *et al.* A New Family of DNA Viruses Causing Disease in Crustaceans from Diverse Aquatic Biomes. **mBio**, v. 11, n. 1, 14 jan. 2020

SUN, C. *et al.* DNA transposons have colonized the genome of the giant virus Pandoravirus salinus. **Bmc Biology**, v. 13, n. 1, 12 jun. 2015.

SUN, S. *et al.* Structural Studies of the Sputnik Virophage. **Journal of Virology**, v. 84, n. 2, p.894-897, 4 nov. 2009.

SUZAN-MONTI, M. *et al.* Ultrastructural Characterization of the Giant Volcano-like Virus Factory of Acanthamoeba polyphaga Mimivirus. **Plos One**, v. 2, n. 3, 28 mar. 2007.

SUZAN-MONTI, M.; LA SCOLA, B.; RAOULT, D. Genomic and evolutionary aspects of Mimivirus. **Virus Research**, v. 117, n. 1, p.145-155, abr. 2006.

THAI, V. *et al.* Structural, Biochemical, and in Vivo Characterization of the First Virally Encoded Cyclophilin from the Mimivirus. **Journal of Molecular Biology**, v. 378, n. 1, p. 71-86, abr. 2008.

- WANG, Jiao *et al.* A novel family of tyrosine integrases encoded by the temperate pleolipovirus SNJ2. **Nucleic Acids Research**, v. 46, n. 5, p. 2521-2536, 18 jan. 2018.
- WICKER, T. *et al.* A unified classification system for eukaryotic transposable elements. **Nature Reviews Genetics**, v. 8, n. 12, p.973-982, dez. 2007.
- WOESE, C. R.; KANDLER, O.; WHEELIS, M. L. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. **Proceedings of the National Academy of Sciences**, v. 87, p. 4576-4579, Jun. 1990
- XIAO, C. *et al.* Cryo-electron Microscopy of the Giant Mimivirus. **Journal of Molecular Biology**, v.353, n.3, p.493-496, out. 2005.
- XIAO, C. *et al.* Structural Studies of the Giant Mimivirus. **Plos Biology**, v. 7, n. 4, p. 959-966, 28 abr. 2009.
- YAU, S. *et al.* Virophage control of antarctic algal host-virus Dynamics. **Proceedings of The National Academy of Sciences**, v. 108, n. 15, p.6163-6168, 28 mar. 2011.
- 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. 4 dez. 2012
- YOSHIKAWA, G. *et al.* Medusavirus, a Novel Large DNA Virus Discovered from Hot Spring Water. **Journal of Virology**, v.93, n.8, 6. fev. 2019
- YUTIN, N. *et al.* A novel group of diverse Polinton-like viruses discovered by metagenome analysis. **Bmc Biology**, v. 13, n. 1, p. 1-2, 11 nov. 2015.
- YUTIN, N. *et al.* Mimiviridae: Clusters of orthologous genes, reconstruction of gene repertoire evolution and proposed expansion of the giant virus family. **Virology Journal**, v. 10, n. 1, p. 1, 04 abr. 2013.
- YUTIN, N.; KAPITONOV, V. K; KOONIN, E. V. A new family of hybrid virophages from an animal gut metagenome. **Biology Direct**, v. 10, n. 1, 25 abr. 2015
- YUTIN, N.; RAOULT, D.; KOONIN, E. V. Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. **Virology Journal**, v. 10, n. 1, 2013.
- YUTIN, N.; RAOULT, D; KOONIN, E. Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. **Virology Journal**, v. 10, n. 1, p. 158, 2013.
- YUTIN, Natalya *et al.* Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. **Virology Journal**, v. 6, n. 1, p. 223, 2009.

ZABLOCKI, O. *et al.* High-Level Diversity of Tailed Phages, Eukaryote-Associated Viruses, and Virophage-Like Elements in the Metaviromes of Antarctic Soils. **Applied and Environmental Microbiology**, v. 80, n. 22, p.6888-6897, 29 ago. 2014.

ZAUBERMAN, N. *et al.* Distinct DNA Exit and Packaging Portals in the Virus *Acanthamoeba polyphaga* mimivirus. **Plos Biology**, v. 6, n. 5, 13 mai. 2008.

ZHANG, X. *et al.* Structure of Sputnik, a virophage, at 3.5-Å resolution. **Proceedings of The National Academy of Sciences**, v. 109, n. 45, p.18431-18436, 22 out. 2012.

ZHOU, J. *et al.* Diversity of Virophages in Metagenomic Data Sets. **Journal of Virology**, v. 87, n. 8, p.4225-4236, 13 fev. 2013.

ZHOU, J. *et al.* Three Novel Virophage Genomes Discovered from Yellowstone Lake Metagenomes. **Journal of Virology**, v. 89, n. 2, p.1278-1285, 12 nov. 2014.

ZIMMERMANN, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. **Journal of Molecular Biology**, v. 430, n. 15, p. 2237-2243, jul. 2018.

ANEXOS

Anexo 1 - Protocolo de preparação do meio de cultura peptona/extrato de levedura/glicose (PYG).

- 1- Os componentes do meio PYG foram pesados na quantidade adequada para o volume de meio a ser produzido:

Componente	Concentração (g/litro)
Sulfato de Magnésio heptahidratado ($MgSO_4 \cdot 7H_2O$) (Merck, Alemanha)	0,98g/litro
Cloreto de cálcio ($CaCl_2$) (Merck, Alemanha)	0,06g/litro
Glicose ($C_6H_{12}O_6$) (Quimibrás, Brasil)	9,0 g/litro
Sulfato de ferro amoniacal hexahidratado ($Fe(NH_4)_2(SO_4) \cdot 6H_2O$) (Merck, Alemanha)	0,02g/litro
Fosfato dibásico de sódio heptahidratado ($Na_2HPO_4 \cdot 7H_2O$) (Merck, Alemanha)	0,40g/litro
Fosfato monobásico de potássio (KH_2PO_4) (Merck, Alemanha)	0,34g/litro
Citrato de sódio dihidratado ($C_6H_5Na_3O_7 \cdot 2H_2O$) (Quimibrás, Brasil)	1,00g/litro
Protease peptona (extrato bactopectona) (Becton Dickinson, E.U.A)	20,0g/litro
Extrato de levedura (Becton Dickinson, E.U.A)	2,0g/litro

- 2- Em seguida, os componentes foram misturados à água destilada e homogeneizados utilizando agitador magnético.
- 3- Após a homogeneização e o ajuste do volume correto, o pH do meio foi ajustado para 6,5, com auxílio do do pHômetro.
- 4- Em seguida, o meio foi filtrado em filtro de café para remoção de restos de materiais não diluídos e, então, encaminhado para a esterilização (autoclave).
- 5- Os meios autoclavados foram submetidos a testes de esterilização e em seguida armazenados em câmara fria até o momento do uso.

Anexo 2 – Best hits obtidos para cada proteína do mimivírus argenteum no BLASTp e no HHpred.

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no HHpred	Prob. %	e-value	Categoria
M1	67	hypothetical protein HIRU_S992 [Hirudovirus strain Sangsue]	100%	2,00E-39	100,00%	sem hit	-	-	Proteína hipotética
M2	467	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	100	2,3e-43	Repetições de domínios com funções diversas
M3	60	hypothetical protein MIMJ_gp0976 [Acanthamoeba polyphaga mimivirus]	100%	4,00E-35	100,00%	Ankyrin repeat containing protein	98,65	0,0000025	Repetições de domínios com funções diversas
M4	210	hypothetical protein MIMJ_gp0975 [Acanthamoeba polyphaga mimivirus]	100%	3,00E-132	100,00%	sem hit	-	-	Proteína hipotética
M5	82	hypothetical protein MIMJ_gp0974 [Acanthamoeba polyphaga mimivirus]	86%	7,00E-40	100,00%	sem hit	-	-	Proteína hipotética
M6	133	hypothetical protein MIMJ_gp0974 [Acanthamoeba polyphaga mimivirus]	99%	5,00E-89	100,00%	sem hit	-	-	Proteína hipotética
M7	210	putative FeOG oxygenase family oxidoreductase [Acanthamoeba polyphaga mimivirus]	100%	2,00E-146	100,00%	ALB homolog 3	99,94	1,2e-24	Enzimas diversas
M8	343	KiLA N-domain protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	DNA binding domain-containing protein	97,96	0,000015	Repetições de domínios com funções diversas
M9	52	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M10	217	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	99%	9,00E-153	100,00%	Ankyrin repeat containing protein	100	3,4e-32	Repetições de domínios com funções diversas
M11	109	hypothetical protein MIMJ_gp0974 [Acanthamoeba polyphaga mimivirus]	100%	4,00E-68	100,00%	sem hit	-	-	Proteína hipotética
M12	62	hypothetical protein MIMJ_gp0973 [Acanthamoeba polyphaga mimivirus]	100%	6,00E-32	100,00%	sem hit	-	-	Proteína hipotética
M13	132	hypothetical protein MIMJ_gp0972 [Acanthamoeba polyphaga mimivirus]	100%	8,00E-93	100,00%	sem hit	-	-	Proteína hipotética
M14	196	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-140	100,00%	sem hit	-	-	Proteína hipotética
M15	579	Acetylcholinesterase [Acanthamoeba castellanii mannavirus]	100%	0,00E-99	99,83%	Acetylcholinesterase-like protein	100	8,2e-82	Enzimas diversas
M16	66	putative membrane protein [Acanthamoeba castellanii mannavirus]	98%	5,00E-36	100,00%	sem hit	-	-	Outras proteínas
M17	89	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	6,00E-55	100,00%	Ankyrin repeat containing protein	99,04	3,3e-9	Repetições de domínios com funções diversas
M18	95	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M19	78	hypothetical protein MIMJ_gp0966 [Acanthamoeba polyphaga mimivirus]	97%	1,00E-36	90,79%	sem hit	-	-	Proteína hipotética
M20	361	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	100	9E-47	Repetições de domínios com funções diversas
M21	110	hypothetical protein HIRU_S15 [Hirudovirus strain Sangsue]	99%	2,00E-70	99,08%	sem hit	-	-	Proteína hipotética
M22	54	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	5,00E-32	100,00%	Chromosomal protein MIC1	99,46	9,8e-14	Outras proteínas
M23	91	hypothetical protein MIMJ_gp0964 [Acanthamoeba polyphaga mimivirus]	100%	3,00E-59	98,99%	sem hit	-	-	Proteína hipotética
M24	161	hypothetical protein MIMJ_gp0963 [Acanthamoeba polyphaga mimivirus]	100%	7,00E-110	100,00%	sem hit	-	-	Proteína hipotética
M25	201	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-141	100,00%	sem hit	-	-	Proteína hipotética
M26	348	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	92%	0,00E-100	100,00%	Ankyrin repeat containing protein	98,63	1,4e-7	Repetições de domínios com funções diversas
M27	321	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	99,97	3,2e-29	Repetições de domínios com funções diversas
M28	698	putative endonuclease/exonuclease/phosphatase [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	Putative endonuclease/exonuclease/phosphatase family protein	99,8	8,7e-17	Enzimas diversas
M29	702	GMC oxidoreductase [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	GMC-type oxidoreductase	100	1,3e-61	Enzimas diversas
M30	269	hypothetical protein MIMJ_gp0957 [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	DUF4579 domain-containing protein (Zn-ribbon)	99,94	2,2e-25	Repetições de domínios com funções diversas
M31	106	hypothetical protein [Acanthamoeba polyphaga mimivirus]	89%	1,00E-57	100,00%	sem hit	-	-	Proteína hipotética
M32	181	hypothetical protein HIRU_S25 [Hirudovirus strain Sangsue]	99%	2,00E-123	99,44%	sem hit	-	-	Proteína hipotética
M33	114	hypothetical protein MIMJ_gp0954 [Acanthamoeba polyphaga mimivirus]	100%	2,00E-76	100,00%	sem hit	-	-	Proteína hipotética
M34	194	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	6,00E-138	99,48%	Ankyrin repeat containing protein	99,9	3,1e-22	Repetições de domínios com funções diversas
M35	271	hypothetical protein MIMJ_gp0953 [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	Aparylyl/Asparaginyl Beta-Hydroxylase	99,97	2,2e-29	Enzimas diversas
M36	169	Transcription factor junoonj [Acanthamoeba castellanii mannavirus]	91%	3,00E-107	100,00%	JunC domain-containing protein	99,17	4,7e-10	Repetições de domínios com funções diversas
M37	429	putative JmC domain-containing protein [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	99,77%	JmC domain-containing protein	99,14	6,8e-10	Repetições de domínios com funções diversas
M38	73	putative ankyrin repeat protein [Hirudovirus strain Sangsue]	98%	4,00E-42	98,61%	Ankyrin repeat containing protein	98,56	0,000005	Repetições de domínios com funções diversas
M39	144	hypothetical protein MIMJ_gp0951 [Acanthamoeba polyphaga mimivirus]	100%	1,00E-97	100,00%	sem hit	-	-	Proteína hipotética
M40	323	putative CDC123-like protein [Acanthamoeba polyphaga mimivirus]	100%	0,00E-100	100,00%	CDC123 protein	100	2,5e-48	Outras proteínas
M41	452	Uncaracterized protein R882 [Acanthamoeba polyphaga mimivirus]	100%	0,00E-118	99,78%	Poly A polymerase; RNA Nucleotidyltransferase, CCA-adding enzyme	98,05	0,00025	Enzimas diversas
M42	207	hypothetical protein [Acanthamoeba polyphaga mimivirus]	99%	1,00E-135	96,60%	Probable serine/threonine-protein kinase	98,24	0,00099	Enzimas diversas
M43	244	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-171	100,00%	Serine/threonine-protein kinase	99,86	3,2e-20	Enzimas diversas
M44	84	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	4,00E-48	100,00%	Kalch repeat and BTB domain-containing protein	97,73	0,00028	Repetições de domínios com funções diversas
M45	187	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	8,00E-135	100,00%	sem hit	-	-	Proteína hipotética
M46	255	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	99,97	1,9e-27	Repetições de domínios com funções diversas
M47	304	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0,00E-100	100,00%	sem hit	-	-	Proteína hipotética
M48	483	KiLA N-terminal domain, N1R.P28 DNA binding protein [Acanthamoeba polyphaga lentivirus]	100%	1,00E-118	100,00%	MLU1-BOX BINDING PROTEIN	98,1	0,000049	Repetições de domínios com funções diversas
M49	171	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9,00E-66	99,11%	Putative outer membrane lipoprotein	99,91	1,3e-21	Outras proteínas
M50	121	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	97,16	0,0011	Repetições de domínios com funções diversas
M51	317	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	sem hit	-	-	Proteína hipotética
M52	169	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-120	100,00%	Ankyrin repeat containing protein	100	2,6e-31	Repetições de domínios com funções diversas
M53	468	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	100	4,2e-38	Repetições de domínios com funções diversas
M54	274	hypothetical protein [Acanthamoeba polyphaga mimivirus]	99%	0,00E-97	80%	P-loop containing nucleoside triphosphate hydrolase	99,72	2,4e-15	Enzimas diversas
M55	76	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4,00E-47	100,00%	sem hit	-	-	Proteína hipotética
M56	575	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5,00E-34	100,00%	MORN repeat-containing protein	99,35	1,00E-11	Repetições de domínios com funções diversas
M57	61	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-141	100,00%	sem hit	-	-	Proteína hipotética
M58	202	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-108	100,00%	sem hit	-	-	Proteína hipotética
M59	163	hypothetical protein Ivs_L763 [Acanthamoeba polyphaga lentivirus]	100%	3,00E-108	100,00%	Ankyrin repeat containing protein	99,84	4,2e-20	Repetições de domínios com funções diversas
M60	426	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,00E-100	100,00%	Ankyrin repeat containing protein	100	1,2e-43	Repetições de domínios com funções diversas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M61	201	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-146	100,00%	Ankyrin repeat containing protein	99,36	4,6E-13	Repetições dominios com funções diversas
M62	65	hypothetical protein HIRU_S54 [Hirudovirus strain Sangsue]	100%	8,00E-32	81,54%	sem hit	-	-	Proteína hipotética
M63	247	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4,00E-178	100,00%	FNIP repeat and Leucine rich repeat protein	91,73	0,37	Repetições dominios com funções diversas
M64	196	BTB super family protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-130	97,45%	BTB POZ domain protein	98,6	8,6E-7	Repetições dominios com funções diversas
M65	677	hypothetical protein HIRU_S57 [Hirudovirus strain Sangsue]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M66	299	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,4E-31	Repetições dominios com funções diversas
M67	528	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	97%	0,0	100,00%	Ankyrin repeat containing protein	100	1,2E-36	Repetições dominios com funções diversas
M68	310	hypothetical protein [Samba virus][APMV]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M69	212	hypothetical protein MIM_M_gp0922 [Acanthamoeba polyphaga minivirus]	100%	2,00E-151	100,00%	sem hit	-	-	Proteína hipotética
M70	106	hypothetical protein MIM_M_gp0921 [Acanthamoeba polyphaga minivirus]	100%	5,00E-70	100,00%	sem hit	-	-	Proteína hipotética
M71	315	hypothetical protein MIM_M_gp0920 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M72	424	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	99,97	3,4E-28	Repetições dominios com funções diversas
M73	180	hypothetical protein MIM_M_gp0918 [Acanthamoeba polyphaga minivirus]	100%	7,00E-130	100,00%	sem hit	-	-	Proteína hipotética
M74	342	putative TPR repeat-containing protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	TPR repeat-containing protein	99,98	4,6E-27	Repetições dominios com funções diversas
M75	123	hypothetical protein Ivs_R749 [Acanthamoeba polyphaga lentivirus]	100%	2,00E-85	100,00%	Deoxyribodipyrimidine photo-lyase-related protein (DPRP)	100	3,5E-34	Reparo de DNA
M76	135	putative transposase [Acanthamoeba polyphaga lentivirus]	96%	6,00E-60	93,24%	sem hit	-	-	Enzimas diversas
M77	129	Unclassified protein R833 [Acanthamoeba polyphaga minivirus]	100%	6,00E-89	100,00%	Deoxyribodipyrimidine photo-lyase-related protein (DPRP)	100	3,7E-35	Reparo de DNA
M78	216	Deoxyribodipyrimidine photolyase-related protein [Acanthamoeba castellanii mannavirus]	99%	1,00E-130	100,00%	Deoxyribodipyrimidine photo-lyase-related protein (DPRP)	99,94	1,1E-24	Enzimas diversas
M79	122	hypothetical protein MIM_M_gp0913 [Acanthamoeba polyphaga minivirus]	81%	2,00E-66	100,00%	sem hit	-	-	Proteína hipotética
M80	513	putative self-like repeat-containing protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	TPR repeat-containing protein	99,92	2,00E-21	Repetições dominios com funções diversas
M81	141	hypothetical protein MIM_M_gp0911 [Acanthamoeba polyphaga minivirus]	100%	3,00E-98	100,00%	sem hit	-	-	Proteína hipotética
M82	519	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,2E-35	Repetições dominios com funções diversas
M83	101	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	2,00E-65	100,00%	Ankyrin repeat containing protein	99,42	2,6E-11	Repetições dominios com funções diversas
M84	545	Putative ankyrin repeat protein R847 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Putative transposase DNA-binding domain	100	1,6E-34	Enzimas diversas
M85	191	putative site-specific integrase-resolvase [Acanthamoeba castellanii mannavirus]	100%	2,00E-133	100,00%	Resolvase-like protein	99,71	9,6E-16	Recombinação
M86	141	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-94	100,00%	Ankyrin repeat containing protein	99,64	4,6E-14	Repetições dominios com funções diversas
M87	437	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,3E-36	Repetições dominios com funções diversas
M88	312	Esterase lipase superfamily protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	alpha beta-Hydrolyases superfamily protein	99,88	7,4E-21	Enzimas diversas
M89	77	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	8,00E-46	100,00%	sem hit	-	-	Proteína hipotética
M90	418	C-terminal divergent BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	97,02%	WD40 repeat-containing protein	99,95	1,4E-25	Repetições dominios com funções diversas
M91	601	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	9,8E-55	Repetições dominios com funções diversas
M92	488	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	Ankyrin repeat containing protein	100	7,8E-46	Repetições dominios com funções diversas
M93	441	hypothetical protein MIM_M_gp0902 [Acanthamoeba polyphaga minivirus]	100%	0,0	99,32%	sem hit	-	-	Proteína hipotética
M94	288	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	0,0	99,65%	Ankyrin repeat containing protein	99,85	4,6E-19	Repetições dominios com funções diversas
M95	626	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	3,00E-50	Repetições dominios com funções diversas
M96	148	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-101	100,00%	GlnB-like superfamily protein	99,69	7,3E-15	Outras proteínas
M97	308	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	5,5E-43	Repetições dominios com funções diversas
M98	102	putative BTB POZ domain-containing protein [Acanthamoeba polyphaga minivirus]	99%	9,00E-65	100,00%	sem hit	-	-	Repetições dominios com funções diversas
M99	544	glucose-methanol-choline oxidoreductase [Acanthamoeba polyphaga minivirus]	100%	0,0	99,82%	GMC oxidoreductase	100	1,7E-48	Enzimas diversas
M100	1611	putative serine threonine-protein kinase receptor [Acanthamoeba polyphaga minivirus]	100%	0,0	99,94%	Serine/threonine-protein kinase	100	1,2E-76	Enzimas diversas
M101	477	putative BTB POZ domain-containing protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	WD40 repeat-containing protein	99,97	5,3E-29	Repetições dominios com funções diversas
M102	433	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M103	78	hypothetical protein MIM_M_gp0892 [Acanthamoeba polyphaga minivirus]	100%	2,00E-31	100,00%	sem hit	-	-	Proteína hipotética
M104	119	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8,00E-79	99,16%	sem hit	-	-	Proteína hipotética
M105	297	hypothetical protein MIM_M_gp0890 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M106	1638	putative serine threonine-protein kinase receptor [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Non-receptor tyrosine-protein kinase	100	1,3E-83	Enzimas diversas
M107	311	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,2E-36	Repetições dominios com funções diversas
M108	221	putative 5'(3')-deoxyribonucleosidase [Acanthamoeba polyphaga minivirus]	100%	2,00E-160	100,00%	putative 5'(3')-deoxyribonucleosidase	99,83	5,7E-19	Enzimas diversas
M109	323	putative proliferating cell nuclear antigen [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	putative proliferating cell nuclear antigen	99,97	1,2E-26	Repetição de DNA
M110	568	hypothetical protein MIM_M_gp0885 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M111	592	putative self-like repeat-containing protein [Acanthamoeba polyphaga minivirus]	87%	0,0	100,00%	putative self-like repeat-containing protein	99,96	1,5E-25	Repetições dominios com funções diversas
M112	305	hypothetical protein MIM_M_gp0877 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	97,37	0,002	Repetições dominios com funções diversas
M113	102	hypothetical protein MIM_M_gp0876 [Acanthamoeba polyphaga minivirus]	100%	2,00E-61	100,00%	sem hit	-	-	Proteína hipotética
M114	990	putative aradine-like ring finger protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	99,83	6,3E-20	Repetições dominios com funções diversas
M115	211	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	1,00E-148	100,00%	Ankyrin repeat containing protein	99,93	3,1E-24	Repetições dominios com funções diversas
M116	229	hypothetical protein MIM_M_gp0872 [Acanthamoeba polyphaga minivirus]	100%	5,00E-166	100,00%	sem hit	-	-	Proteína hipotética
M117	169	cytochrome P450 family protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-119	100,00%	cytochrome P450 family protein	99,79	4,5E-18	Enzimas diversas
M118	546	probable lanosterol 14-alpha demethylase [Acanthamoeba polyphaga minivirus]	99%	0,0	99,82%	Cytochrome P450	100	8,6E-33	Enzimas diversas
M119	447	probable 7-dehydrocholesterol reductase [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Delta(14)-sterol reductase	100	1,7E-65	Enzimas diversas
M120	117	hypothetical protein [Acanthamoeba castellanii mannavirus]	92%	2,00E-73	100,00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hlppred	Prob. %	e-value	Categoria
M121	192	hypothetical protein MIM1_gp0865 [Acanthamoeba polyphaga minimivirus]	100%	4,00E-136	100,00%	sem hit	-	-	Proteína hipotética
M122	65	hypothetical protein MIM1_gp0867 [Acanthamoeba polyphaga minimivirus]	81%	9,02E-27	100,00%	sem hit	-	-	Proteína hipotética
M123	140	hypothetical protein MIM1_gp0867 [Acanthamoeba polyphaga minimivirus]	96%	1,02E-88	100,00%	sem hit	-	-	Proteína hipotética
M124	243	putative metal-dependent phosphohydrolase, HD domain [APMV]	100%	0,0	100,00%	sem hit	99,92	5,1E-23	Enzimas diversas
M125	247	putative purine phosphorylase [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Purine and uridine phosphorylases superfamily	99,95	2,7E-25	Enzimas diversas
M126	131	hypothetical protein MIM1_gp0864 [Acanthamoeba polyphaga minimivirus]	100%	2,02E-88	100,00%	sem hit	-	-	Proteína hipotética
M127	114	hypothetical protein MIM1_gp0863 [Acanthamoeba polyphaga minimivirus]	100%	2,02E-75	100,00%	sem hit	-	-	Proteína hipotética
M128	201	hypothetical protein [Acanthamoeba castellanii mannavirus]	96%	1,00E-136	100,00%	sem hit	-	-	Proteína hipotética
M129	269	hypothetical protein Ivs_L700 [Acanthamoeba polyphaga lentivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	8,7E-38	Repetições dominios com funções diversas
M130	445	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M131	142	hypothetical protein MIM1_gp0859 [Acanthamoeba polyphaga minimivirus]	100%	3,00E-101	100,00%	sem hit	99,94	5E-26	Enzimas diversas
M132	1297	RING-finger-containing E3 ubiquitin ligase [Acanthamoeba polyphaga minimivirus]	100%	0,0	99,92%	Ubiquitin conjugating enzyme	-	-	Enzimas diversas
M133	378	hypothetical protein MIM1_gp0884 [Acanthamoeba polyphaga minimivirus]	99%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M134	120	hypothetical protein MIM1_gp0883 [Acanthamoeba polyphaga minimivirus]	100%	4,00E-84	100,00%	sem hit	-	-	Proteína hipotética
M135	438	hypothetical protein MIM1_R819 [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	90,48	1	Repetições dominios com funções diversas
M136	1651	putative serine/threonine-protein kinase receptor [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	N-receptor tyrosine-protein kinase	100	8,3E-84	Enzimas diversas
M137	153	hypothetical protein [Samba virus]	99%	2,00E-104	99,34%	sem hit	-	-	Proteína hipotética
M138	198	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-142	100,00%	sem hit	-	-	Proteína hipotética
M139	311	hypothetical protein MIM1_R784 [Acanthamoeba polyphaga minimivirus]	100%	0,0	99,68%	Ankyrin repeat containing protein	100	2,3E-33	Repetições dominios com funções diversas
M140	332	DNA primase small subunit [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	DNA primase small subunit	99,94	1,1E-25	Repetição de DNA
M141	224	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9,00E-163	100,00%	sem hit	-	-	Proteína hipotética
M142	228	hypothetical protein MIM1_gp0854 [Acanthamoeba polyphaga minimivirus]	100%	2,00E-165	100,00%	sem hit	-	-	Proteína hipotética
M143	304	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	94%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M144	118	hypothetical protein HIRU_S132 [Hirudovirus strain Sangueu]	100%	2,00E-78	100,00%	Ankyrin repeat containing protein	100	2,8E-34	Repetições dominios com funções diversas
M145	422	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M146	498	WD40 family protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,3E-48	Repetições dominios com funções diversas
M147	450	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	97%	0,0	100,00%	WD40 repeat-containing protein	99,95	1,9E-25	Repetições dominios com funções diversas
M148	492	putative BTB/POZ domain and WD-repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	99,80%	Ankyrin repeat containing protein	100	1,00E-45	Repetições dominios com funções diversas
M149	108	hypothetical protein MIM1_gp0847 [Acanthamoeba polyphaga minimivirus]	100%	5,00E-69	100,00%	sem hit	-	-	Proteína hipotética
M150	118	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	98%	1,00E-69	96,53%	F-box domain e LRR containing protein	90,21	0,45	Repetições dominios com funções diversas
M151	524	WD40 family protein [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	WD40 repeat-containing protein	99,97	4,3E-27	Repetições dominios com funções diversas
M152	126	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-87	100,00%	Nucleotidyltransferase superfamily	94,79	0,068	Enzimas diversas
M153	151	hypothetical protein MIM1_gp0843 [Acanthamoeba polyphaga minimivirus]	100%	6,00E-108	100,00%	sem hit	-	-	Proteína hipotética
M154	289	uncharacterized dTDP-4-dehydrothiamose reductase domain protein [APMV]	100%	0,0	100,00%	UDP-L-Rhamnose Synthase	100	1,00E-37	Enzimas diversas
M155	103	hypothetical protein MIM1_gp0841 [Acanthamoeba polyphaga minimivirus]	100%	6,00E-68	100,00%	sem hit	-	-	Proteína hipotética
M156	257	hypothetical protein MIM1_gp0840 [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	MIR domain superfamily	97,94	0,0011	Proteína hipotética
M157	410	ankyrin-containing protein [Hirudovirus strain Sangueu]	100%	0,0	99,76%	Ankyrin repeat containing protein	100	6,8E-46	Repetições dominios com funções diversas
M158	77	putative chaperon protein DnaJ [Acanthamoeba polyphaga minimivirus]	98%	1,00E-46	100,00%	Cysteine-rich domain of the chaperone protein DnaJ	98,18	0,000054	Chaperona
M159	342	putative AAA-type ATPase [Acanthamoeba polyphaga minimivirus]	100%	7,00E-147	100,00%	AAA-type ATPase	100	1,6E-40	Enzimas diversas
M160	203	putative PAN domain-containing protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	PAN domain-containing protein	99,47	6,4E-12	Repetições dominios com funções diversas
M161	633	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	6,7E-44	Repetições dominios com funções diversas
M162	489	putative BTB/POZ domain-containing protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	WD40 repeat-containing protein	99,96	6,7E-25	Repetições dominios com funções diversas
M163	295	hypothetical protein MIM1_gp0833 [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	4,3E-32	Repetições dominios com funções diversas
M164	538	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	99%	0,0	100,00%	Ankyrin repeat containing protein	100	4,7E-46	Repetições dominios com funções diversas
M165	626	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	1,7E-51	Repetições dominios com funções diversas
M166	86	hypothetical protein MIM1_gp0830 [Acanthamoeba polyphaga minimivirus]	89%	8,00E-49	100,00%	sem hit	-	-	Proteína hipotética
M167	281	hypothetical protein MIM1_L767b [Acanthamoeba polyphaga minimivirus]	92%	0,0	99,62%	Ankyrin repeat containing protein	99,92	7,00E-24	Repetições dominios com funções diversas
M168	215	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	8,00E-143	99,53%	Ankyrin repeat containing protein	99,89	3,00E-21	Repetições dominios com funções diversas
M169	194	hypothetical protein MIM1_L766b [Acanthamoeba polyphaga minimivirus]	95%	1,00E-124	99,46%	Ankyrin repeat containing protein	99,39	8,1E-12	Repetições dominios com funções diversas
M170	83	hypothetical protein MIM1_L766a [Acanthamoeba polyphaga minimivirus]	98%	3,00E-43	99,90%	Ankyrin repeat containing protein	99,62	0,79	Repetições dominios com funções diversas
M171	292	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	99,66%	Ankyrin repeat containing protein	100	8,1E-39	Repetições dominios com funções diversas
M172	473	putative BTB/POZ domain-containing protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	WD40 repeat-containing protein	99,97	6,1E-27	Repetições dominios com funções diversas
M173	138	hypothetical protein MIM1_L764b [Acanthamoeba polyphaga minimivirus]	100%	2,00E-90	100,00%	sem hit	-	-	Proteína hipotética
M174	185	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-120	100,00%	Ankyrin repeat containing protein	96,36	0,16	Repetições dominios com funções diversas
M175	147	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	7,00E-100	100,00%	Ankyrin repeat containing protein	99	4,1E-9	Repetições dominios com funções diversas
M176	262	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DUF4379 domain-containing protein	99,82	1,7E-19	Proteína hipotética
M177	251	hypothetical protein MIM1_gp0822 [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M178	159	hypothetical protein MIM1_gp0821 [Acanthamoeba polyphaga minimivirus]	100%	3,00E-111	100,00%	sem hit	-	-	Proteína hipotética
M179	501	putative ankyrin repeat protein [Acanthamoeba polyphaga minimivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	4,8E-43	Repetições dominios com funções diversas
M180	214	hypothetical protein MIM1_gp0819 [Acanthamoeba polyphaga minimivirus]	100%	2,00E-155	99,53%	Putative 3-methyladenine DNA glycosylase	100	8,00E-35	Reparo de DNA

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrpred	Prob. %	e-value	Categoria
M181	78	hypothetical protein [Acanthamoeba castellanii mimivirus]	100%	1.00E-48	100.00%	sem hit	-	-	Proteína hipotética
M182	183	putative 5'(3')-deoxyribonucleotidase [Acanthamoeba polyphaga mimivirus]	100%	3.00E-130	100.00%	putative 5'(3')-deoxyribonucleotidase	99.67	4.2e-15	Enzimas diversas
M183	396	putative F-box protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Poly A. polymerase head domain-containing protein	97.38	0.012	Transcrição
M184	180	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-128	100.00%	LRR-containing protein	99.21	3.4e-10	Repetições/dominios com funções diversas
M185	298	putative Fe-S-cluster redox enzyme [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Radical SAM superfamily protein	99.95	6.00E-25	Enzimas diversas
M186	160	hypothetical protein MIM1_gp0815 [Acanthamoeba polyphaga mimivirus]	100%	1.00E-86	100.00%	sem hit	-	-	Proteína hipotética
M187	139	hypothetical protein HIRU_S170 [Hirudovirus strain Saugseu]	97%	1.00E-75	100.00%	sem hit	-	-	Proteína hipotética
M188	74	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	1.00E-35	100.00%	sem hit	-	-	Proteína hipotética
M189	179	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	6.00E-119	99.44%	sem hit	-	-	Proteína hipotética
M190	86	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-52	100.00%	sem hit	-	-	Proteína hipotética
M191	751	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	F-box/LRR repeat - containing protein	99.57	6.9e-13	Repetições/dominios com funções diversas
M192	215	hypothetical protein MIM1_gp0811 [Acanthamoeba polyphaga mimivirus]	100%	7.00E-160	100.00%	sem hit	-	-	Proteína hipotética
M193	149	hypothetical protein MIM1_gp0810 [Acanthamoeba polyphaga mimivirus]	100%	3.00E-100	100.00%	sem hit	-	-	Proteína hipotética
M194	338	hypothetical protein MIM1_gp0809 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M195	224	putative homeobox protein [Acanthamoeba polyphaga mimivirus]	100%	1.00E-155	100.00%	putative homeobox protein	98.56	4.6e-7	Transcrição
M196	1058	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-90	100.00%	Ankyrin repeat containing protein	100	1.9e-49	Repetições/dominios com funções diversas
M197	215	hypothetical protein MIM1_gp0805 [Acanthamoeba polyphaga mimivirus]	100%	4.00E-158	100.00%	sem hit	-	-	Proteína hipotética
M198	138	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-90	100.00%	sem hit	-	-	Proteína hipotética
M199	192	Uncharacterized protein R744 [Acanthamoeba polyphaga mimivirus]	90%	1.00E-115	94.48%	sem hit	-	-	Proteína hipotética
M200	211	hypothetical protein [Acanthamoeba castellanii mannavirus]	91%	6.00E-137	100.00%	sem hit	-	-	Proteína hipotética
M201	358	putative BTB/POZ domain-containing protein [Acanthamoeba polyphaga mimivirus]	95%	0.0	100.00%	POZ domain-containing protein	99.51	1.5e-13	Repetições/dominios com funções diversas
M202	182	hypothetical protein MIM1_gp0800 [Acanthamoeba polyphaga mimivirus]	100%	1.00E-126	100.00%	sem hit	-	-	Proteína hipotética
M203	187	putative mitochondrial protein [Acanthamoeba polyphaga mimivirus]	100%	1.00E-139	100.00%	sem hit	-	-	Proteína hipotética
M204	490	WD40 family protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-60	100.00%	WD40 repeat-containing protein	100	7.1e-30	Repetições/dominios com funções diversas
M205	107	WD40 family protein [Acanthamoeba castellanii mannavirus]	93%	2.00E-60	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M206	399	WD40 family protein [Acanthamoeba castellanii mannavirus]	98%	0.0	100.00%	WD40 repeat-containing protein	99.92	1.4e-23	Repetições/dominios com funções diversas
M207	248	hypothetical protein MIM1_gp0796 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M208	276	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M209	446	hypothetical protein MIM1_gp0794 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	9.4e-35	Repetições/dominios com funções diversas
M210	652	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	2E-47	Repetições/dominios com funções diversas
M211	532	ABC transporter ATP-binding domain protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	ABC transporter ATP-binding domain protein	100	5.9e-59	Repetições/dominios com funções diversas
M212	148	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	2.00E-105	99.32%	sem hit	-	-	Proteína hipotética
M213	518	putative BTB/POZ domain and WD-repeat protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	WD40 repeat-containing protein	99.97	7E-28	Repetições/dominios com funções diversas
M214	406	putative CtrQ-like protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	AAA family ATPase	99.85	4.5e-18	Enzimas diversas
M215	189	hypothetical protein HIRU_S197 [Hirudovirus strain Saugseu]	100%	1.00E-134	100.00%	sem hit	-	-	Proteína hipotética
M216	572	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0.0	100.00%	RNA-binding protein	99.97	3.3e-27	Repetições/dominios com funções diversas
M217	284	hypothetical protein MIM1_gp0786 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	phosphomethylhydrazide kinase family protein	99.42	6.6e-12	Enzimas diversas
M218	252	peptide chain release factor eRF1 [Samba virus]	99%	1.00E-178	99.60%	Peptide chain release factor eRF1	100	1.5e-30	Tradução
M219	70	peptide chain release factor 1 N [Hirudovirus strain Saugseu]	100%	1.00E-40	98.55%	N-terminal domain of eukaryotic peptide chain release factor subunit 1, ERF1	99.74	3E-18	Tradução
M220	224	hypothetical protein MIM1_gp0784 [Acanthamoeba polyphaga mimivirus]	48%	9.00E-14	100.00%	Eukaryotic peptide chain release factor subunit 1	93.45	0.095	Tradução
M221	224	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-154	100.00%	SEA Cysteine rich and Ictin C-type domain-containing protein	89.34	0.52	Repetições/dominios com funções diversas
M222	218	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	7.00E-158	100.00%	sem hit	-	-	Proteína hipotética
M223	221	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Enzimas diversas
M224	872	putative bifunctional metalloprotease ubiquitin-protein ligase [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Putative bifunctional metalloprotease	100	2.1e-31	Enzimas diversas
M225	270	putative chemotaxis protein ChdD [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Putative chemotaxis protein ChdD	100	6.1e-34	Outras proteínas
M226	274	endomucase VIII-like protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Endomucase VIII-like protein	100	2.5e-55	Reparo de DNA
M227	349	hypothetical protein Ivs_L626 [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	PAN domain-containing protein	99.38	1.6e-11	Repetições/dominios com funções diversas
M228	130	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-91	100.00%	sem hit	-	-	Proteína hipotética
M229	206	hypothetical protein MIM1_gp0776 [Acanthamoeba polyphaga mimivirus]	100%	2.00E-146	100.00%	sem hit	-	-	Proteína hipotética
M230	171	putative glutamine amidotransferase-like protein [Acanthamoeba polyphaga mimivirus]	100%	3.00E-118	100.00%	Putative glutamine amidotransferase-like protein	99.9	1.8e-21	Enzimas diversas
M231	91	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	4.00E-40	87.78%	sem hit	-	-	Proteína hipotética
M232	54	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	5.00E-25	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M233	189	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	2.00E-126	100.00%	Ankyrin repeat containing protein	99.92	1.1e-23	Repetições/dominios com funções diversas
M234	115	hypothetical protein MIM1_gp0773 [Acanthamoeba polyphaga mimivirus]	100%	2.00E-78	100.00%	sem hit	-	-	Proteína hipotética
M235	214	putative membrane protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-145	100.00%	sem hit	-	-	Outras proteínas
M236	452	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	93.96%	sem hit	-	-	Outras proteínas
M237	376	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0.0	100.00%	sem hit	-	-	Outras proteínas
M238	530	hypothetical protein MIM1_gp0770 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Phage lysozyme domain-containing protein	96.96	0.022	Outras proteínas
M239	334	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	98.20%	AAA family ATPase	99.76	6.6e-17	Enzimas diversas
M240	194	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-144	100.00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M241	388	probable ubiquitin-conjugating enzyme E2 [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	Ubiquitin-conjugating enzyme E2	99.88	1.1e-21	Enzimas diversas
M242	202	uncharacterized phosphoglycerate mutase family protein [Acanthamoeba polyphaga minivirus]	100%	2.00E-145	100.00%	Phosphoglycerate mutase-like superfamily	99.97	2.1e-28	Enzimas diversas
M243	281	putative glycosyltransferase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative glycosyltransferase	99.97	3.3e-29	Enzimas diversas
M244	420	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M245	313	hypothetical protein MIM_I_gp0763 [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M246	463	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0.0	100.00%	sem hit	-	-	Enzimas diversas
M247	108	hypothetical protein vs_R611 [Acanthamoeba polyphaga lentivirius]	100%	2.00E-70	100.00%	sem hit	96.93	0.1	Enzimas diversas
M248	135	hypothetical protein MIM_I_gp0761 [Acanthamoeba polyphaga minivirus]	100%	1.00E-93	100.00%	sem hit	-	-	Proteína hipotética
M249	495	putative protein kinase [Acanthamoeba castellanii mannavirus]	99%	0.0	98.38%	putative protein kinase	99.66	5.1e-15	Enzimas diversas
M250	240	putative protein kinase [Acanthamoeba castellanii mannavirus]	100%	3.00E-163	100.00%	sem hit	-	-	Enzimas diversas
M251	203	putative protein kinase [Acanthamoeba castellanii mannavirus]	100%	3.00E-147	100.00%	sem hit	-	-	Enzimas diversas
M252	401	Serpin [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Serpin	100	5E-52	Enzimas diversas
M253	455	putative procollagen-lysine-2-oxoglutarate dioxygenase [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	Putative procollagen-lysine-2-oxoglutarate dioxygenase	100	9E-36	Enzimas diversas
M254	300	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M255	284	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-167	100.00%	sem hit	-	-	Proteína hipotética
M256	129	hypothetical protein vs_L603 [Acanthamoeba polyphaga lentivirius]	99%	3.00E-84	100.00%	Ankyrin repeat containing protein	99.84	3.4e-19	Repetições/dominios com funções diversas
M257	63	hypothetical protein MIM_I_gp0754 [Acanthamoeba polyphaga minivirus]	100%	2.00E-37	100.00%	sem hit	-	-	Proteína hipotética
M258	313	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.95	5.4e-26	Repetições/dominios com funções diversas
M259	144	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-99	100.00%	sem hit	-	-	Proteína hipotética
M260	128	putative potassium channel protein [Hirudovirus strain Sangsue]	100%	7.00E-84	98.43%	Putative Potassium channel protein	99.03	1.6e-7	Outras proteínas
M261	149	probable methylated-DNA-protein-cysteine methyltransferase [Acanthamoeba polyphaga minivirus]	100%	9.00E-106	100.00%	Putative methylated-DNA-protein-cysteine methyltransferase	100	7.5e-34	Reparo de DNA
M262	354	hypothetical protein MIM_I_gp0749 [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M263	275	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M264	374	hypothetical protein MIM_I_gp0746 [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	NHL repeat-containing protein	99.69	3.4e-13	Repetições/dominios com funções diversas
M265	255	probable UDP-N-acetylglucosamine pyrophosphorylase [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	UDP-N-acetylglucosamine pyrophosphorylase	99.07	6.7e-28	Enzimas diversas
M266	222	hypothetical protein MIM_I_gp0744 [Acanthamoeba polyphaga minivirus]	100%	3.00E-153	100.00%	TNF-like superfamily protein	99.07	1.2e-8	Outras proteínas
M267	330	putative UV damage endonuclease [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	putative UV damage endonuclease (UvrE)	100	5.5e-34	Reparo de DNA
M268	219	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-153	100.00%	sem hit	-	-	Proteína hipotética
M269	72	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M270	150	hypothetical protein HIRU_S244 [Hirudovirus strain Sangsue]	100%	2.00E-98	99.33%	sem hit	-	-	Proteína hipotética
M271	333	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M272	192	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9.00E-135	100.00%	sem hit	-	-	Proteína hipotética
M273	476	Endonuclease Exonuclease phosphatase family protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Endonuclease Exonuclease phosphatase family	99.78	5.9e-17	Enzimas diversas
M274	196	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9.00E-141	100.00%	sem hit	-	-	Proteína hipotética
M275	114	hypothetical protein MIM_I_gp0736 [Acanthamoeba polyphaga minivirus]	100%	3.00E-76	100.00%	Proteína com domínio transmembrana	95.95	0.92	Outras proteínas
M276	151	hypothetical protein MIM_I_gp0734 [Acanthamoeba polyphaga minivirus]	100%	1.00E-103	100.00%	sem hit	-	-	Proteína hipotética
M277	70	hypothetical protein MIM_I_gp0733 [Acanthamoeba polyphaga minivirus]	100%	5.00E-40	100.00%	sem hit	-	-	Proteína hipotética
M278	66	putative serine-threonine-protein kinase [Acanthamoeba polyphaga minivirus]	95%	2.00E-31	98.41%	sem hit	-	-	Enzimas diversas
M279	600	putative serine-threonine-protein kinase [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	putative serine-threonine-protein kinase	99.98	3.9e-31	Enzimas diversas
M280	225	putative SET domain-containing protein [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	putative SET domain-containing protein	99.93	3.4e-24	Repetições/dominios com funções diversas
M281	84	hypothetical protein MIM_I_gp0729 [Acanthamoeba polyphaga minivirus]	71%	3.00E-33	100.00%	sem hit	-	-	Proteína hipotética
M282	102	hypothetical protein MIM_I_gp0728 [Acanthamoeba polyphaga minivirus]	100%	3.00E-68	100.00%	sem hit	-	-	Proteína hipotética
M283	472	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	6.2e-38	Repetições/dominios com funções diversas
M284	134	hypothetical protein MIM_I_L674b [Acanthamoeba polyphaga minivirus]	100%	3.00E-89	100.00%	POZ domain-containing protein	99.54	5E-14	Repetições/dominios com funções diversas
M285	348	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M286	60	No significant similarity found.	-	-	-	sem hit	-	-	Proteína hipotética
M287	63	hypothetical protein MIM_I_gp0725 [Acanthamoeba polyphaga minivirus]	100%	4.00E-33	98.41%	sem hit	-	-	Proteína hipotética
M288	96	hypothetical protein [Acanthamoeba polyphaga minivirus]	100%	2.00E-60	100.00%	sem hit	-	-	Proteína hipotética
M289	530	hypothetical protein [Acanthamoeba polyphaga minivirus]	100%	0.0	96.89%	sem hit	-	-	Proteína hipotética
M290	111	hypothetical protein MIM_I_gp0723 [Acanthamoeba polyphaga minivirus]	100%	3.00E-79	100.00%	sem hit	99.98	9.00E-30	Enzimas diversas
M291	154	hypothetical protein MIM_I_gp0722 [Acanthamoeba polyphaga minivirus]	98%	6.00E-106	100.00%	sem hit	-	-	Proteína hipotética
M292	539	serine-threonine kinase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	putative serine-threonine-protein kinase	100	3E-31	Enzimas diversas
M293	1932	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	99%	0.0	100.00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M294	1291	collagen triple helix repeat containing protein [Hirudovirus strain Sangsue]	87%	0.0	98.46%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M295	148	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6.00E-103	99.32%	Zinc-ribbon domain-containing protein	99.46	2.3e-12	Repetições/dominios com funções diversas
M296	59	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-32	96.61%	sem hit	-	-	Proteína hipotética
M297	198	putative ATPase [Acanthamoeba castellanii mannavirus]	100%	4.00E-140	100.00%	ATP-dependent DNA helicase	99.45	5.4e-13	Enzimas diversas
M298	152	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	2.00E-102	100.00%	Ankyrin repeat containing protein	99.85	4.7e-19	Repetições/dominios com funções diversas
M299	467	putative oxidoreductase [Acanthamoeba polyphaga minivirus]	100%	0.0	100.00%	Putative oxidoreductase	99.93	6.9e-24	Enzimas diversas
M300	408	Arginyl-RNA synthetase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Arginyl-RNA synthetase	100	1.2e-48	Tradução

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrped	Prob. %	e-value	Categoria
M301	148	arginyl-tRNA synthetase [Acanthamoeba polyphaga minivirius]	79%	2.00E-74	98.31%	Argnyl-tRNA synthetase	84.27	0.88	Tradução
M302	462	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M303	73	hypothetical protein HIRU_S272 [Hirudovirus strain Sangsue]	100%	3.00E-40	100.00%	sem hit	-	-	Proteína hipotética
M304	211	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-148	100.00%	sem hit	-	-	Proteína hipotética
M305	206	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-148	100.00%	sem hit	-	-	Proteína hipotética
M306	508	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	EGF-like domain-containing protein	97.98	0.00038	Repetições/dominios com funções diversas
M307	156	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-106	100.00%	sem hit	-	-	Proteína hipotética
M308	371	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative phosphoglycerate mutase-like superfamily protein	99.82	7.2e-20	Enzimas diversas
M309	204	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	7.00E-142	100.00%	AAA family ATPase	96.9	0.0025	Enzimas diversas
M310	479	FAD dependent oxidoreductase family protein [Acanthamoeba polyphaga minivirius]	100%	0.0	100.00%	FAD dependent oxidoreductase family protein	100	4.4e-30	Enzimas diversas
M311	324	putative glycosyltransferase [Acanthamoeba polyphaga minivirius]	100%	2.00E-163	100.00%	Putative glycosyltransferase	99.93	1.3e-27	Enzimas diversas
M312	224	hypothetical protein MIM1_R654b [Acanthamoeba polyphaga minivirius]	100%	2.00E-144	100.00%	Glycosyltransferase family 10 protein	100	2.9e-36	Enzimas diversas
M313	159	glycosyltransferase family 10 [Acanthamoeba castellanii mannavirus]	96%	3.00E-101	99.34%	sem hit	-	-	Enzimas diversas
M314	181	glycosyltransferase family 10 [Acanthamoeba castellanii mannavirus]	100%	2.00E-122	100.00%	Putative glycosyltransferase	99.96	5.2e-28	Enzimas diversas
M315	227	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-168	100.00%	Cysteine proteinases	98.29	0.000088	Enzimas diversas
M316	348	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0.0	99.71%	POZ domain-containing protein	93.09	0.58	Repetições/dominios com funções diversas
M317	328	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M318	106	hypothetical protein [Acanthamoeba castellanii mannavirus]	83%	1.00E-55	100.00%	sem hit	-	-	Proteína hipotética
M319	146	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-97	100.00%	sem hit	-	-	Proteína hipotética
M320	119	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-76	100.00%	sem hit	-	-	Proteína hipotética
M321	77	hypothetical protein MIM1_gp0699 [Acanthamoeba polyphaga minivirius]	100%	2.00E-48	100.00%	sem hit	-	-	Proteína hipotética
M322	281	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	P-loop containing nucleotide triphosphate hydrolase	99.46	1.5e-11	Enzimas diversas
M323	400	NHL repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	NHL repeat-containing protein	99.6	7.3e-12	Repetições/dominios com funções diversas
M324	185	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-123	100.00%	sem hit	-	-	Proteína hipotética
M325	273	hypothetical protein MIM1_gp0695 [Acanthamoeba polyphaga minivirius]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M326	407	low complexity hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M327	143	phosphatidylinositol-binding protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-93	95.10%	Phosphatidylinositol-binding protein	100	7.0E-32	Outras proteínas
M328	359	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M329	526	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	NHL repeat-containing protein	99.34	8.3e-11	Repetições/dominios com funções diversas
M330	461	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	95.84%	SET and RING associated domain	93.51	0.42	Repetições/dominios com funções diversas
M331	550	Methionyl-tRNA synthetase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Methionyl-tRNA synthetase	100	1.2e-72	Tradução
M332	134	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-91	100.00%	FNIP repeat-containing protein	95.5	0.011	Repetições/dominios com funções diversas
M333	153	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6.00E-107	100.00%	FNIP repeat-containing protein	91.38	0.25	Repetições/dominios com funções diversas
M334	152	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-107	98.68%	sem hit	-	-	Proteína hipotética
M335	81	hypothetical protein [Samba virus]	100%	4.00E-46	100.00%	FNIP repeat-containing protein	96.11	0.19	Repetições/dominios com funções diversas
M336	315	FNIP repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	FNIP repeat-containing protein	99.28	8.6e-9	Repetições/dominios com funções diversas
M337	70	hypothetical protein MIM1_gp0684 [Acanthamoeba polyphaga minivirius]	100%	1.00E-42	100.00%	BTP/POZ and Zinc-ribbon domain-containing protein	99.36	1.1e-11	Repetições/dominios com funções diversas
M338	408	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Zinc-ribbon domain-containing protein	99.89	8.1e-22	Repetições/dominios com funções diversas
M339	378	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat-containing protein	100	7.8e-47	Repetições/dominios com funções diversas
M340	111	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-70	100.00%	Ankyrin repeat-containing protein	99.02	4.9e-9	Repetições/dominios com funções diversas
M341	59	No significant similarity found.	-	-	-	sem hit	-	-	Proteína hipotética
M342	216	Bubb-type mannosase-specific lectin [Acanthamoeba castellanii mannavirus]	92%	2.00E-139	100.00%	alpha-D-mannosase-specific plant lectins	99.03	8.2e-8	Outras proteínas
M343	53	putative Fe-S oxidoreductase [Samba virus]	84%	2.00E-10	82.22%	sem hit	-	-	Enzimas diversas
M344	321	putative N-acylsphingosine amidohydrolase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative N-acylethanolamine-hydrolyzing acid amidase	100	1.6e-44	Enzimas diversas
M345	227	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	7.00E-163	100.00%	sem hit	-	-	Proteína hipotética
M346	200	uncharacterized transmembrane protein [Acanthamoeba polyphaga minivirius]	100%	7.00E-144	100.00%	Cytochrome b5-like Heme Steroid binding domain	98.08	0.00018	Enzimas diversas
M347	322	hypothetical protein MIM1_gp0676 [Acanthamoeba polyphaga minivirius]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M348	643	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M349	599	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	vWA-like superfamily protein	98.26	0.000038	Outras proteínas
M350	167	GTP-binding protein-like protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-116	100.00%	sem hit	-	-	Outras proteínas
M351	479	GTP binding elongation factor eF-Tu [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	GTP binding elongation factor eF-Tu	100	2E-42	Tradução
M352	269	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	TNF-like family protein	98.03	0.00025	Outras proteínas
M353	212	putative tyrosine-protein phosphatase [Acanthamoeba polyphaga minivirius]	100%	5.00E-156	100.00%	putative tyrosine-protein phosphatase	99.78	5.7e-17	Enzimas diversas
M354	358	putative N-myristoyltransferase [Hirudovirus strain Sangsue]	100%	0.0	100.00%	putative N-myristoyltransferase	100	4E-47	Enzimas diversas
M355	368	Patinin phospholipase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Patinin phospholipase	99.96	1.6e-26	Enzimas diversas
M356	606	putative glucosamine-fructose-6-phosphate aminotransferase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative glucosamine-fructose-6-phosphate aminotransferase	100	1.2e-56	Enzimas diversas
M357	161	putative GTPase [Acanthamoeba castellanii mannavirus]	100%	3.00E-103	99.38%	sem hit	-	-	Enzimas diversas
M358	61	cytochrome b5-like protein [Hirudovirus strain Sangsue]	98%	9.00E-29	88.33%	Cytochrome b5-like protein	99.07	1.7e-8	Enzimas diversas
M359	392	putative GTPase [Hirudovirus strain Sangsue]	100%	0.0	98.00%	putative GTPase	99.83	6.6e-18	Enzimas diversas
M360	170	hypothetical protein MIM1_gp0665 [Acanthamoeba polyphaga minivirius]	100%	8.00E-125	100.00%	YuA-A-like family protein	100	3.6e-34	Outras proteínas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrp2d	Prob. %	e-value	Categoria
M361	353	hypothetical protein MIM1_gp0662 [Acanthamoeba polyphaga mnavivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M362	701	Phosphoinositide 3-kinase [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Phosphoinositide 3-kinase	100	5.1e-75	Enzimas diversas
M363	303	Stomatin family protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Stomatin family protein	99.9	1.8e-20	Outras diversas
M364	392	hypothetical protein [Acanthamoeba castellani mannavirus]	98%	0.0	100.00%	JmjC domain-containing histone demethylation protein	99.85	1.4e-20	Repetições/dominios com funções diversas
M365	160	Mannose-6P isomerase [Acanthamoeba castellani mannavirus]	100%	2.00E-115	100.00%	Mannose-6P isomerase	99.43	6.5e-11	Outras proteínas
M366	953	P-loop ATPase GTPase-containing domain protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	P-loop ATPase GTPase-containing domain protein	97.81	0.000088	Enzimas diversas
M367	316	Proline rich protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	CI-2 family of serine protease inhibitors (serpin)	97.24	0.00038	Outras proteínas
M368	149	hypothetical protein [Acanthamoeba castellani mannavirus]	79%	6.00E-68	92.68%	sem hit	-	-	Proteína hipotética
M369	103	hypothetical protein [Acanthamoeba castellani mannavirus]	96%	3.00E-61	98.99%	sem hit	-	-	Proteína hipotética
M370	86	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	4.00E-51	100.00%	sem hit	-	-	Proteína hipotética
M371	88	hypothetical protein MIM1_gp0654 [Acanthamoeba polyphaga mnavivirus]	100%	5.00E-56	100.00%	sem hit	-	-	Proteína hipotética
M372	150	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	1.00E-101	99.33%	sem hit	-	-	Proteína hipotética
M373	225	hypothetical protein HIRU_S331 [Hirudovirus strain Sangsue]	99%	2.00E-159	98.21%	sem hit	-	-	Proteína hipotética
M374	170	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	2.00E-119	100.00%	sem hit	-	-	Proteína hipotética
M375	234	Cyclophilin family protein [Acanthamoeba castellani mannavirus]	100%	1.00E-170	100.00%	Cyclophilin type peptidyl-prolyl cis-trans isomerase CLD	99.91	2.7e-21	Chaperona
M376	95	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	2.00E-63	100.00%	sem hit	-	-	Proteína hipotética
M377	375	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M378	263	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	98%	0.0	100.00%	Ankyrin repeat containing protein	99.97	1.6e-29	Repetições/dominios com funções diversas
M379	68	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	73%	1.00E-23	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M380	304	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	3.7e-40	Repetições/dominios com funções diversas
M381	303	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.97	5.4e-28	Repetições/dominios com funções diversas
M382	315	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.85	1.8e-25	Repetições/dominios com funções diversas
M383	359	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	99.72%	Ankyrin repeat containing protein	99.98	2.4e-31	Repetições/dominios com funções diversas
M384	237	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	1.00E-164	100.00%	Ankyrin repeat containing protein	99.96	1.8e-27	Repetições/dominios com funções diversas
M385	53	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	2.00E-27	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M386	276	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.97	1.9e-29	Repetições/dominios com funções diversas
M387	292	Thiol oxidoreductase E1OR [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Probable FAD-linked sulfhydryl oxidase	100	7.4e-59	Enzimas diversas
M388	70	hypothetical protein MIM1_gp0639 [Acanthamoeba polyphaga mnavivirus]	100%	8.00E-45	100.00%	RING finger domain, (Zinc finger, C3HC4 type)	98.41	7.7e-7	Repetições/dominios com funções diversas
M389	277	putative phospholipase carboxylesterase family protein [Hirudovirus strain Sangsue]	100%	0.0	100.00%	alfa beta-hydroxylases	99.96	5.6e-26	Enzimas diversas
M390	399	hypothetical protein MIM1_L594 [Acanthamoeba polyphaga mnavivirus]	100%	0.0	99.75%	Probable ATP-dependent RNA helicase	99.97	3.5e-30	Enzimas diversas
M391	242	putative prolyl 4-hydroxylase [Acanthamoeba polyphaga mnavivirus]	100%	4.00E-180	100.00%	Putative prolyl 4-hydroxylase	99.93	1E-23	Enzimas diversas
M392	126	RING-finger-containing E3 ubiquitin ligase [Acanthamoeba castellani mannavirus]	100%	4.00E-67	100.00%	sem hit	-	-	Enzimas diversas
M393	720	RING-finger-containing E3 ubiquitin ligase [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Helicase conserved C-terminal and RING finger domain	100	1.7e-50	Enzimas diversas
M394	238	hypothetical protein MIM1_gp0634 [Acanthamoeba polyphaga mnavivirus]	99%	1.00E-171	99.58%	Transmembrane protein	95.27	0.14	Outras proteínas
M395	190	hypothetical protein MIM1_L590 [Acanthamoeba polyphaga mnavivirus]	100%	4.00E-139	100.00%	sem hit	-	-	Proteína hipotética
M396	272	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	8.6e-31	Repetições/dominios com funções diversas
M397	397	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M398	163	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	6.00E-117	100.00%	sem hit	-	-	Proteína hipotética
M399	103	hypothetical protein MIM1_gp0629 [Acanthamoeba polyphaga mnavivirus]	100%	1.00E-66	100.00%	sem hit	-	-	Proteína hipotética
M400	107	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	2.00E-68	100.00%	sem hit	-	-	Proteína hipotética
M401	411	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Acetyltransferase (GNAT) family	99.43	3.1e-11	Enzimas diversas
M402	104	hypothetical protein MIM1_gp0636 [Acanthamoeba polyphaga mnavivirus]	100%	2.00E-67	100.00%	Mis12-Nr1f1 homolog protein	94.86	0.54	Outras proteínas
M403	179	hypothetical protein MIM1_gp0625 [Acanthamoeba polyphaga mnavivirus]	100%	2.00E-121	100.00%	sem hit	-	-	Proteína hipotética
M404	741	ATP-dependent helicase HrpA [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	ATP-dependent helicase HrpA	100	2.2e-64	Enzimas diversas
M405	473	ankyrin repeat-containing protein [Acanthamoeba castellani mannavirus]	100%	0.0	99.37%	Ankyrin repeat containing protein	100	9.2e-42	Repetições/dominios com funções diversas
M406	463	putative ankyrin repeat protein [Acanthamoeba castellani mannavirus]	97%	0.0	98.83%	Ankyrin repeat containing protein	100	5.5e-43	Repetições/dominios com funções diversas
M407	438	hypothetical protein [Acanthamoeba polyphaga mnavivirus]	100%	8.00E-68	100.00%	Ankyrin repeat containing protein	100	2.1e-41	Repetições/dominios com funções diversas
M408	104	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M409	296	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M410	285	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M411	245	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	5.00E-175	100.00%	sem hit	-	-	Proteína hipotética
M412	219	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	4.00E-155	100.00%	sem hit	-	-	Proteína hipotética
M413	485	AAA family ATPase [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	AAA family ATPase	100	3.6e-34	Enzimas diversas
M414	297	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Patatin-like phospholipase	100	2.9e-32	Enzimas diversas
M415	262	hypothetical protein [Acanthamoeba castellani mannavirus]	94%	3.00E-175	100.00%	sem hit	-	-	Proteína hipotética
M416	1041	UvDREP helicase family protein [Acanthamoeba castellani mannavirus]	99%	0.0	100.00%	UvDREP helicase family protein	100	3.5e-33	Enzimas diversas
M417	320	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	0.0	100.00%	Metalloprotein-like family protein	100	1.7e-48	Outras proteínas
M418	53	hypothetical protein [Acanthamoeba polyphaga mnavivirus]	67%	1.00E-11	91.67%	sem hit	-	-	Proteína hipotética
M419	88	hypothetical protein MIM1_gp0610 [Acanthamoeba polyphaga mnavivirus]	98%	5.00E-54	97.70%	sem hit	-	-	Proteína hipotética
M420	126	hypothetical protein [Acanthamoeba castellani mannavirus]	100%	1.00E-85	100.00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hlppred	Prob. %	e-value	Categoria
M421	76	hypothetical protein [Acanthamoeba polyphaga mimivirus]	94%	5,0E-36	91,67%	sem hit	-	-	Proteína hipotética
M422	144	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9,0E-103	100,00%	sem hit	-	-	Proteína hipotética
M423	370	glutamine synthetase [Acanthamoeba castellanii mannavirus]	95%	0,0	100,00%	Glutamine synthetase	100	2E-67	Enzimas diversas
M424	137	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6,0E-95	100,00%	sem hit	-	-	Proteína hipotética
M425	573	putative ATP-dependent RNA helicase [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	helicase conserved C-terminal domain-containing protein	100	1,7e-40	Enzimas diversas
M426	321	hypothetical protein MIM1_gp0605 [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M427	236	hypothetical protein MIM1_gp0604 [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M428	114	Nuomod4 XRN endonuclease [Acanthamoeba castellanii mannavirus]	96%	7,0E-71	95,43%	NUMOD4 motif HNH endonuclease	98,65	5,2e-8	Enzimas diversas
M429	78	nuomod4 XRN endonuclease [Acanthamoeba polyphaga mimivirus]	94%	8,0E-44	98,63%	DNA-binding domain	98,62	0,0000023	Repetições/dominios com funções diversas
M430	226	Helix-turn-helix XRE-family like protein [Acanthamoeba castellanii mannavirus]	100%	2,0E-163	100,00%	Helix-turn-helix XRE-family like protein	98,41	0,000004	Outras proteínas
M431	68	putative membrane protein [Acanthamoeba castellanii mannavirus]	100%	2,0E-37	100,00%	sem hit	-	-	Outras proteínas
M432	465	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M433	160	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	7,0E-114	100,00%	sem hit	-	-	Proteína hipotética
M434	278	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M435	34	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M436	325	putative DNA repair protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	SMC domain-containing protein	99,8	1,0E-15	Repetições/dominios com funções diversas
M437	227	putative DNA repair protein [Acanthamoeba polyphaga mimivirus]	100%	7,0E-142	98,67%	sem hit	-	-	Reparo de DNA
M438	749	putative DNA repair protein [Acanthamoeba castellanii mannavirus]	99%	0,0	99,73%	Putative DNA repair protein	99,94	6,8e-24	Reparo de DNA
M439	65	hypothetical protein [Samba virus]	95%	3,0E-33	100,00%	sem hit	-	-	Enzimas diversas
M440	141	Thioredoxin [Acanthamoeba castellanii mannavirus]	100%	8,0E-98	100,00%	Thioredoxin	99,79	1E-16	Proteína hipotética
M441	122	hypothetical protein lvs_R474 [Acanthamoeba polyphaga lentilevirus]	100%	4,0E-85	100,00%	sem hit	-	-	Enzimas diversas
M442	213	MAMA_R546 [Acanthamoeba castellanii mannavirus]	100%	7,0E-143	100,00%	sem hit	-	-	Proteína hipotética
M443	282	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M444	132	transcription initiation factor TFIIIB [Acanthamoeba polyphaga mimivirus]	99%	3,0E-85	99,24%	HMG-box family protein	96,07	0,054	Repetições/dominios com funções diversas
M445	409	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	Putative Nucleotidyltransferase	100	4,9e-48	Enzimas diversas
M446	518	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Aukym repeat containing protein	97,33	0,0023	Repetições/dominios com funções diversas
M447	308	hypothetical protein HIRU_S392 [Hirudovirus strain Saugse]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M448	285	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Putative endonuclease	97,35	0,00028	Enzimas diversas
M449	167	hypothetical protein MIM1_gp0592 [Acanthamoeba polyphaga mimivirus]	100%	8,0E-103	100,00%	Aukym repeat containing protein	99,96	1,6e-27	Repetições/dominios com funções diversas
M450	277	hypothetical protein lvs_R476 [Acanthamoeba polyphaga lentilevirus]	100%	1,0E-163	100,00%	sem hit	-	-	Proteína hipotética
M451	234	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M452	327	ADP-ribosyltransferase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Putative Nucleotidyltransferase	100	3E-91	Enzimas diversas
M453	363	hypothetical protein lvs_R470 [Acanthamoeba polyphaga lentilevirus]	100%	0,0	100,00%	ADP-ribosyltransferase	100	1,5e-37	Enzimas diversas
M454	81	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Leucine-rich repeat-containing protein	99,91	4,8e-22	Repetições/dominios com funções diversas
M455	330	putative potassium channel voltage dependent tetramerization domain protein [APNAV]	93%	9,0E-41	100,00%	Leucine-rich repeat-containing protein	97,24	0,00089	Repetições/dominios com funções diversas
M456	169	C-terminal divergent BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	97,87%	sem hit	-	-	Outras proteínas
M457	695	VV18 helicase [Acanthamoeba polyphaga lentilevirus]	94%	5,0E-111	100,00%	C-terminal divergent BTB domain-containing protein	99,62	2,6e-15	Repetições/dominios com funções diversas
M458	299	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Pre-mRNA-splicing factor ATP-dependent RNA helicase	100	7,7e-85	Enzimas diversas
M459	74	putative membrane protein [Acanthamoeba castellanii mannavirus]	100%	6,0E-46	100,00%	Restriction endonuclease-like	94,72	0,23	Enzimas diversas
M460	60	hypothetical protein MIM1_gp0579 [Acanthamoeba polyphaga mimivirus]	100%	5,0E-32	100,00%	sem hit	-	-	Outras proteínas
M461	445	putative ATP-dependent RNA helicase [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	Putative ATP-dependent RNA helicase	-	-	Proteína hipotética
M462	144	hypothetical protein lvs_L464 [Acanthamoeba polyphaga lentilevirus]	100%	1,0E-95	100,00%	DNA primase small subunit	100	3,6e-46	Enzimas diversas
M463	108	hypothetical protein lvs_L464 [Acanthamoeba polyphaga lentilevirus]	100%	3,0E-69	100,00%	sem hit	-	-	Repetição de DNA
M464	123	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,0E-77	100,00%	sem hit	-	-	Proteína hipotética
M465	169	Uncharacterized protein R335 [Acanthamoeba polyphaga mimivirus]	100%	7,0E-119	100,00%	sem hit	-	-	Proteína hipotética
M466	205	hypothetical protein lvs_R454 [Acanthamoeba polyphaga lentilevirus]	97%	1,0E-81	100,00%	sem hit	-	-	Proteína hipotética
M468	133	hypothetical protein MIM1_L534 [Acanthamoeba polyphaga mimivirus]	100%	3,0E-91	100,00%	sem hit	-	-	Proteína hipotética
M469	124	mudx hydrolase [Acanthamoeba castellanii mannavirus]	100%	2,0E-81	100,00%	mRNA decapping enzyme-like	98,79	3,6e-8	Proteína hipotética
M470	508	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Transcrição
M471	468	Cytochrome P450 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DnaQ-like 3'-5' exonuclease	99,79	1,9e-17	Enzimas diversas
M472	298	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	Cytochrome P450	100	1,4e-55	Enzimas diversas
M473	494	putative DNA helicase [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M474	164	hypothetical protein MIM1_gp0569 [Acanthamoeba polyphaga mimivirus]	100%	2,0E-115	100,00%	Putative ATP-dependent DNA helicase	100	2,6e-35	Enzimas diversas
M475	785	XRN 5'-3' exonuclease [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Eukaryotic translation initiation factor 4E	100	3,6e-35	Enzimas diversas
M476	345	putative triacylglycerol lipase [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	XRN 5'-3' exonuclease	100	8,8e-113	Tradução
M477	297	hypothetical protein lvs_R454 [Acanthamoeba polyphaga lentilevirus]	100%	0,0	100,00%	sem hit	-	-	Enzimas diversas
M478	94	hypothetical protein lvs_R454 [Acanthamoeba polyphaga lentilevirus]	100%	1,0E-51	100,00%	Putative alpha beta-hydrolase	100	1,3e-31	Enzimas diversas
M479	166	hypothetical protein MIM1_R335 [Acanthamoeba polyphaga mimivirus]	100%	8,0E-108	100,00%	sem hit	-	-	Proteína hipotética
M480	392	MdAT-like NTP pyrophosphohydrolase [Samba virus]	100%	0,0	99,49%	Nudix superfamily protein	95,98	0,01	Enzimas diversas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M481	163	Integral membrane protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-110	100,00%	Rhomboid family transmembrane protein	99,78	4,2e-17	Outras proteínas
M482	264	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M483	1134	ubiquitin-conjugating enzyme E2 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ubiquitin-conjugating enzyme	100	1,4e-44	Enzimas diversas
M484	321	ubiquitin-conjugating enzyme E2 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Enzimas diversas
M485	846	BTB family protein [Acanthamoeba castellanii mannavirus]	97%	0,0	99,40%	POZ domain-containing protein	99,41	5,7e-13	Repetições/dominios com funções diversas
M486	330	putative peptidase [Acanthamoeba castellanii mannavirus]	100%	0,0	90,88%	Putative peptidase	100	2,2e-47	Enzimas diversas
M487	102	putative zinc metalloprotease [Acanthamoeba polyphaga minivirus]	89%	1,00E-46	92,22%	sem hit	-	-	Enzimas diversas
M488	110	putative peptidase [Acanthamoeba polyphaga minivirus]	100%	4,00E-71	100,00%	sem hit	-	-	Enzimas diversas
M489	112	unknown [Acanthamoeba polyphaga minivirus]	100%	9,00E-70	99,10%	sem hit	-	-	Proteína hipotética
M490	74	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5,00E-38	100,00%	sem hit	-	-	Proteína hipotética
M491	265	hypothetical protein MIM1_gp0555 [Acanthamoeba polyphaga minivirus]	98%	3,00E-178	99,24%	sem hit	-	-	Enzimas diversas
M492	301	Serine Threonine protein kinase [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	Serine/Threonine protein kinase	100	2E-44	Enzimas diversas
M493	221	hypothetical protein MIM1_gp0557 [Acanthamoeba polyphaga minivirus]	100%	9,00E-153	100,00%	sem hit	-	-	Proteína hipotética
M494	579	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	99,83%	sem hit	-	-	Proteína hipotética
M495	345	hypothetical protein MIM1_gp0555 [Acanthamoeba polyphaga minivirus]	98%	0,0	99,70%	sem hit	-	-	Proteína hipotética
M496	168	hypothetical protein MIM1_gp0554 [Acanthamoeba polyphaga minivirus]	100%	8,00E-92	82,86%	sem hit	-	-	Proteína hipotética
M497	74	hypothetical protein [Samba virus]	100%	3,00E-36	100,00%	sem hit	-	-	Proteína hipotética
M498	83	hypothetical protein MIM1_gp0553 [Acanthamoeba polyphaga minivirus]	100%	1,00E-54	100,00%	sem hit	-	-	Proteína hipotética
M499	168	hypothetical protein MIM1_gp0552 [Acanthamoeba polyphaga minivirus]	100%	5,00E-115	100,00%	sem hit	-	-	Proteína hipotética
M500	193	hypothetical protein MIM1_gp0552 [Acanthamoeba polyphaga minivirus]	100%	3,00E-54	100,00%	sem hit	-	-	Proteína hipotética
M501	309	Deoxynucleotide monophosphate kinase [Acanthamoeba castellanii mannavirus]	100%	5,00E-143	100,00%	Deoxynucleotide monophosphate kinase	99,88	3,4e-20	Enzimas diversas
M502	467	ftsJ methyltransferase domain-containing protein 1 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	FtsJ-like methyltransferase	99,9	5,4e-22	Enzimas diversas
M503	363	ftsJ methyltransferase domain-containing protein 1 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	FtsJ-like methyltransferase	99,97	3E-28	Enzimas diversas
M504	292	putative replication factor C small subunit [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Putative replication factor C small subunit	100	1,5e-31	Replicação de DNA
M505	335	hypothetical protein MIM1_L509 [Acanthamoeba polyphaga minivirus]	100%	0,0	99,66%	Calcineurin-like phosphoesterase domain	99,91	8,5e-22	Outras proteínas
M506	65	hypothetical protein MIM1_gp0546 [Acanthamoeba polyphaga minivirus]	100%	3,00E-35	100,00%	sem hit	-	-	Repetições/dominios com funções diversas
M507	82	hypothetical protein MIM1_gp0546 [Acanthamoeba polyphaga minivirus]	100%	5,00E-48	100,00%	sem hit	-	-	Proteína hipotética
M508	405	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M509	312	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M510	92	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-54	100,00%	sem hit	-	-	Proteína hipotética
M511	232	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-165	100,00%	sem hit	-	-	Proteína hipotética
M512	512	DNA-directed RNA polymerase subunit 1 [Acanthamoeba polyphaga minivirus]	100%	2,00E-37	100,00%	DNA-directed RNA polymerase subunit 1	100	1,6e-71	Transcrição
M513	66	hypothetical protein MIM1_gp0546 [Acanthamoeba polyphaga minivirus]	99%	0,0	100,00%	Nucleic acid-binding protein	96,91	0,13	Proteína hipotética
M514	323	hypothetical protein MIM1_gp0545 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	DNA-directed RNA polymerase subunit 1	100	1,2e-77	Outras proteínas
M515	440	DNA-directed RNA polymerase subunit 1 [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Group I intron endonuclease	99,61	2,3e-14	Transcrição
M516	256	Group I intron endonuclease [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DNA-directed RNA polymerase subunit 1	100	6,4e-93	Enzimas diversas
M517	514	DNA directed RNA polymerase (D) subunit 1 [Acanthamoeba castellanii mannavirus]	100%	7,00E-52	100,00%	sem hit	-	-	Transcrição
M518	83	hypothetical protein HIRU_S447 [Hirudovirus strain Sangseu]	100%	1,00E-130	100,00%	sem hit	-	-	Proteína hipotética
M519	183	hypothetical protein [Acanthamoeba polyphaga minivirus]	100%	1,00E-70	100,00%	sem hit	-	-	Proteína hipotética
M520	110	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Bacterial fluorinating enzyme, C-terminal domain	90,72	0,31	Outras proteínas
M521	344	putative replication factor C small subunit [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Putative replication factor C small subunit	100	6,1e-28	Replicação de DNA
M522	422	probable zinc-type alcohol dehydrogenase-like protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Putative zinc-type alcohol dehydrogenase-like protein	100	7,2e-29	Enzimas diversas
M523	563	bifunctional dithyrotolate reductase-thymidylate synthase [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Bifunctional dithyrotolate reductase-thymidylate synthase	100	8,6e-87	Enzimas diversas
M524	148	Translation initiation factor 4E [Acanthamoeba castellanii mannavirus]	98%	2,00E-102	100,00%	Translation initiation factor 4E	99,97	3,5e-30	Tradução
M525	114	Translation initiation factor 4E [Acanthamoeba castellanii mannavirus]	100%	6,00E-74	100,00%	sem hit	-	-	Tradução
M526	140	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5,00E-95	100,00%	sem hit	-	-	Proteína hipotética
M527	273	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M528	293	probable DNA polymerase sliding clamp [Acanthamoeba polyphaga minivirus]	100%	0,0	99,66%	Proliferating cell nuclear antigen	100	1,3e-28	Replicação de DNA
M529	66	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M530	179	hypothetical protein MIM1_gp0531 [Acanthamoeba polyphaga minivirus]	100%	7,00E-129	100,00%	Non-histone chromosomal protein MC1	96,66	0,013	Outras proteínas
M531	239	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-166	96,76%	sem hit	-	-	Proteína hipotética
M532	304	hypothetical protein MIM1_gp0529 [Acanthamoeba polyphaga minivirus]	100%	0,0	98,63%	sem hit	-	-	Proteína hipotética
M533	82	hypothetical protein MIM1_gp0528 [Acanthamoeba polyphaga minivirus]	100%	3,00E-51	100,00%	sem hit	-	-	Proteína hipotética
M534	198	hypothetical protein MIM1_gp0528 [Acanthamoeba polyphaga minivirus]	100%	8,00E-144	100,00%	Capsid protein	100	2,6e-34	Proteínas estruturais/morfogêneses
M535	559	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	98,3	0,0000044	Repetições/dominios com funções diversas
M536	93	hypothetical protein MIM1_gp0525 [Acanthamoeba polyphaga minivirus]	100%	4,00E-58	100,00%	sem hit	-	-	Proteína hipotética
M537	335	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	98,51%	sem hit	-	-	Proteína hipotética
M538	223	putative PAN domain-containing protein [Acanthamoeba polyphaga minivirus]	100%	4,00E-163	100,00%	Putative PAN domain-containing protein	99,12	5,7e-9	Repetições/dominios com funções diversas
M539	170	hypothetical protein MIM1_gp0521 [Acanthamoeba polyphaga minivirus]	100%	4,00E-122	100,00%	DNA-directed RNA polymerase subunit P	96,26	0,0051	Transcrição
M540	1719	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	Ankyrin repeat containing protein	98,79	1,5e-8	Repetições/dominios com funções diversas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hlppred	Prob. %	e-value	Categoria
M601	88	hypothetical protein MIM1_gp0466[Acanthamoeba polyphaga mimivirus]	100%	8,00E-56	100,00%	sem hit	-	-	Proteína hipotética
M602	407	hypothetical protein MIM1_gp0465[Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M603	139	hypothetical protein MIM1_gp0464[Acanthamoeba polyphaga mimivirus]	100%	7,00E-111	100,00%	sem hit	-	-	Proteína hipotética
M604	76	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	1,00E-47	100,00%	sem hit	-	-	Proteína hipotética
M605	57	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	5,00E-32	100,00%	sem hit	-	-	Proteína hipotética
M606	192	putative HD phosphohydrolase [Acanthamoeba castellanii mannavirus]	100%	9,00E-138	100,00%	Putative HD phosphohydrolase	98,79	0,000016	Enzimas diversas
M607	186	hypothetical protein MIM1_R431b [Acanthamoeba polyphaga mimivirus]	100%	6,00E-131	100,00%	HRDC and 3'-5' exonuclease domain	99,24	1,7E-10	Enzimas diversas
M608	212	3'-5' exonuclease [Acanthamoeba castellanii mannavirus]	95%	2,00E-137	100,00%	Putative 3'-5' exonuclease	99,48	5,8E-13	Enzimas diversas
M609	187	putative late transcription factor [Acanthamoeba polyphaga mimivirus]	83%	2,00E-100	100,00%	sem hit	-	-	Transcrição
M610	379	putative transcription factor [Acanthamoeba polyphaga mimivirus]	99%	0,0	100,00%	putative transcription factor	95,73	0,36	Transcrição
M611	194	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6,00E-140	100,00%	sem hit	-	-	Proteína hipotética
M612	120	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-80	100,00%	sem hit	-	-	Proteína hipotética
M613	581	capsid protein D13L [Hirudovirus strain Sangsue]	100%	0,0	100,00%	Capsid protein	100	3,3E-72	Proteínas estruturais/morfogêneses
M614	132	hypothetical protein MIM1_R424 [Acanthamoeba polyphaga mimivirus]	100%	2,00E-87	100,00%	sem hit	-	-	Proteína hipotética
M615	193	hypothetical protein MIM1_gp0454[Acanthamoeba polyphaga mimivirus]	97%	3,00E-132	100,00%	NUMOD4 motif HNH homing endonuclease	99,04	1,6E-9	Enzimas diversas
M616	530	Zn finger-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Putative HNH restriction endonuclease	99,31	1,8E-11	Enzimas diversas
M617	232	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6,00E-168	100,00%	sem hit	-	-	Proteína hipotética
M618	209	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-150	100,00%	sem hit	-	-	Proteína hipotética
M619	244	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6,00E-175	100,00%	sem hit	-	-	Proteína hipotética
M620	342	uncharacterized short-chain type dehydrogenase/reductase [APN(V)]	100%	0,0	100,00%	Uncharacterized short-chain type dehydrogenase/reductase	89,61	0,63	Repetições/dominios com funções diversas
M621	137	nucleoside diphosphate kinase [Acanthamoeba polyphaga mimivirus]	100%	2,00E-97	100,00%	Nucleoside diphosphate kinase	99,91	8,6E-22	Enzimas diversas
M622	491	hypothetical protein MIM1_gp0447[Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Enzimas diversas
M623	105	hypothetical protein MIM1_gp0446[Acanthamoeba polyphaga mimivirus]	100%	5,00E-67	100,00%	sem hit	-	-	Proteína hipotética
M624	415	FNIP repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	FNIP repeat-containing protein	99,44	1,7E-9	Repetições/dominios com funções diversas
M625	75	putative F-box and FNIP repeat-containing protein [Hirudovirus strain Sangsue]	93%	2,00E-40	94,29%	FNIP repeat-containing protein	96,72	0,033	Repetições/dominios com funções diversas
M626	336	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M627	306	hypothetical protein [Acanthamoeba castellanii mannavirus]	93%	0,0	100,00%	COCH zinc finger family protein	97,06	0,00047	Repetições/dominios com funções diversas
M628	246	replication factor C large subunit [Acanthamoeba castellanii mannavirus]	100%	1,00E-176	100,00%	replication factor C large subunit	100	1,4E-34	Replicação de DNA
M629	235	replication factor C large subunit [Acanthamoeba castellanii mannavirus]	92%	1,00E-157	100,00%	replication factor C large subunit	99,73	7,4E-16	Replicação de DNA
M630	661	putative core protein [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteínas estruturais/morfogêneses
M631	185	hypothetical protein [Acanthamoeba polyphaga mimivirus]	100%	3,00E-127	100,00%	Capsid protein	100	1,3E-44	Enzimas diversas
M632	223	putative papain-like cysteine peptidase [Acanthamoeba castellanii mannavirus]	100%	2,00E-160	100,00%	sem hit	-	-	Enzimas diversas
M633	365	putative RNA methyltransferase [Acanthamoeba polyphaga mimivirus]	97%	0,0	100,00%	putative RNA methyltransferase	100	1,8E-51	Enzimas diversas
M634	191	Alkylated DNA repair protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-141	100,00%	Alkylated DNA repair protein	99,97	9,8E-29	Reparo de DNA
M635	338	FtsI-like methyltransferase family protein [Acanthamoeba castellanii mannavirus]	100%	0,0	98,07%	Putative RNA methyltransferase	100	3,3E-30	Enzimas diversas
M636	263	Esterase lipase superfamily protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Esterase lipase superfamily protein	99,95	7,8E-26	Enzimas diversas
M637	734	hypothetical protein MIM1_gp0433[Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M638	798	Low complexity hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	putative serine threonine-protein kinase	100	4,6E-32	Enzimas diversas
M639	131	hypothetical protein MIM1_gp0431[Acanthamoeba polyphaga mimivirus]	99%	4,00E-88	100,00%	Protein serine-threonine phosphatase	99,93	4,9E-24	Proteína hipotética
M640	467	putative serine threonine-protein kinase [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	ADP-ribosylation superfamily protein	97,64	0,00049	Enzimas diversas
M641	332	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M642	362	hypothetical protein MIM1_R398 [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	Protein-tyrosine-phosphatase	99,95	2,7E-25	Enzimas diversas
M643	574	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M644	519	DNA or RNA helicase of superfamily II [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	ATP-dependent DNA helicase	99,94	4,2E-27	Enzimas diversas
M645	186	replication factor C small subunit [Acanthamoeba castellanii mannavirus]	100%	5,00E-126	100,00%	Replication factor C small subunit	99,87	7,8E-19	Replicação de DNA
M646	133	putative replication factor C small subunit [Acanthamoeba polyphaga mimivirus]	100%	3,00E-90	100,00%	Replication factor C small subunit	99,74	1,4E-15	Replicação de DNA
M647	448	putative HD domain-containing protein [Acanthamoeba polyphaga mimivirus]	100%	0,0	100,00%	Putative HD domain-containing protein	100	2,3E-80	Enzimas diversas
M648	634	HSP70 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	HSP70	100	4,6E-69	Chaperona
M649	1163	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0,0	99,86%	ADP-ribosylation superfamily protein	97,64	0,00049	Outras proteínas
M650	1083	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Polyphosphate kinase C-terminal domain family protein	99,59	5,8E-14	Proteína hipotética
M651	268	Endonuclease exonuclease-phosphatase family protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Leucine-rich repeat-containing protein	99,94	3,2E-24	Enzimas diversas
M652	265	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Bromodomain-containing protein	98,03	0,000092	Proteína hipotética
M653	473	putative endonuclease of the XPG family [Acanthamoeba castellanii mannavirus]	100%	3,00E-96	100,00%	Putative endonuclease of the XPG family	100	5,3E-44	Enzimas diversas
M654	141	hypothetical protein MIM1_gp0416[Acanthamoeba castellanii mannavirus]	100%	3,00E-96	100,00%	sem hit	-	-	Proteína hipotética
M655	132	hypothetical protein MIM1_gp0415 [Acanthamoeba polyphaga mimivirus]	100%	2,00E-89	100,00%	sem hit	-	-	Proteína hipotética
M656	348	divergent methyltransferase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	FtsI-like methyltransferase	99,87	6,9E-20	Enzimas diversas
M657	1121	probable mRNA-capping enzyme [Acanthamoeba polyphaga mimivirus]	99%	0,0	99,91%	Putative mRNA-capping enzyme	100	5,3E-33	Transcrição
M658	57	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M659	233	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-168	100,00%	Polyphosphate kinase C-terminal domain family protein	99,59	5,8E-14	Enzimas diversas
M660	78	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-43	100,00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrp2	Prob. %	e-value	Categoria
M661	338	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	99.70%	Leucine-rich repeat-containing protein	99.94	1.7e-23	Repetições/dominos com funções diversas
M662	153	hypothetical protein MIMI_gp0410 [Acanthamoeba polyphaga manamvirus]	99%	4.00E-104	100.00%	sem hit	-	-	Proteína hipotética
M663	131	hypothetical protein MIMI_gp0410 [Acanthamoeba polyphaga manamvirus]	100%	4.00E-89	100.00%	sem hit	-	-	Proteína hipotética
M664	140	hypothetical protein MIMI_gp0409 [Acanthamoeba polyphaga manamvirus]	88%	9.00E-81	100.00%	Bromodomain-containing protein	97.96	0.00017	Repetições/dominos com funções diversas
M665	1116	putative NTPase [Acanthamoeba polyphaga manamvirus]	99%	3.00E-138	100.00%	Putative NTPase	100	2.8e-38	Enzimas diversas
M666	199	putative DNA-directed RNA polymerase subunit [Acanthamoeba polyphaga manamvirus]	100%	3.00E-138	100.00%	putative DNA-directed RNA polymerase subunit	99.89	6.2e-21	Transcrição
M667	360	putative diphosphonitric polyphosphate phosphohydrolyase [APMV]	100%	0.0	100.00%	Nudix superfamily protein	98.27	0.0000034	Enzimas diversas
M668	225	putative N-acetylglucosaminyl phosphatidylinositol deacetylase [APMV]	100%	8.00E-163	100.00%	putative N-acetylglucosaminyl phosphatidylinositol deacetylase	99.94	1.5e-25	Outras proteínas
M669	251	putative glycosyltransferase [Acanthamoeba polyphaga manamvirus]	100%	0.0	99.60%	Putative glycosyltransferase	99.78	5.2e-18	Enzimas diversas
M670	250	hypothetical protein MIMI_gp0404 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Putative Cysteine peptidase	99.85	2.00E-19	Enzimas diversas
M671	705	ankyrin repeat-containing protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	7.3e-43	Repetições/dominos com funções diversas
M672	144	hypothetical protein MIMI_gp0404 [Acanthamoeba polyphaga manamvirus]	100%	3.00E-98	100.00%	sem hit	-	-	Proteína hipotética
M673	445	hypothetical protein MIMI_gp0401 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Replication factor A protein	97.79	0.00045	Replicação de DNA
M674	143	probable FAD-linked sulfhydryl oxidase [Acanthamoeba polyphaga manamvirus]	100%	2.00E-100	100.00%	FAD-linked sulfhydryl oxidase domain	99.72	4.8e-19	Enzimas diversas
M675	483	hypothetical protein MIMI_gp0398 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	PDZ and N-terminal domain of MutM-like DNA repair proteins	100	1.4e-36	Reparo de DNA
M676	1654	hypothetical protein vs_R302 [Acanthamoeba polyphaga lentivirus]	88%	0.0	99.46%	Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16	100	1.9e-66	Enzimas diversas
M677	236	hypothetical protein HIRU_S592 [Hirudovirus strain Sangsue]	100%	2.00E-157	98.72%	Ankyrin repeat containing protein	94.83	0.063	Repetições/dominos com funções diversas
M678	65	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	1.00E-56	100.00%	sem hit	-	-	Proteína hipotética
M679	534	DEAD SNF2 helicase [Acanthamoeba castellanii manamvirus]	99%	0.0	100.00%	DEAD SNF2 helicase	100	1.9e-39	Enzimas diversas
M680	245	putative glycosyltransferase [Acanthamoeba castellanii manamvirus]	100%	1.00E-180	100.00%	Putative glycosyltransferase	99.95	4.3e-26	Enzimas diversas
M681	346	Thioredoxin [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	Thioredoxin-like protein	99.47	2.1e-12	Enzimas diversas
M682	262	hypothetical protein MIMI_gp0391 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M683	151	hypothetical protein [Acanthamoeba castellanii manamvirus]	95%	1.00E-98	100.00%	sem hit	-	-	Proteína hipotética
M684	1107	putative DNA mismatch repair protein MutS-like protein [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Putative DNA mismatch repair protein MutS-like protein	100	3.6e-105	Reparo de DNA
M685	287	hypothetical protein MIMI_gp0388 [Acanthamoeba polyphaga manamvirus]	97%	0.0	100.00%	DNA BETA-GLUCOSYL TRANSFERASE	99.86	9.00E-20	Enzimas diversas
M686	73	putative DNA-directed RNA polymerase II subunit N [Acanthamoeba polyphaga manamvirus]	100%	2.00E-46	100.00%	putative DNA-directed RNA polymerase II subunit N	99.98	9.7e-33	Transcrição
M687	481	Low complexity hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	96.98%	Capsid protein	97.2	0.005	Proteínas estruturais/morfogêneses
M688	228	hypothetical protein MIMI_gp0385 [Acanthamoeba polyphaga manamvirus]	96%	4.00E-145	100.00%	sem hit	-	-	Proteína hipotética
M689	347	hypothetical protein vs_L291 [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	Capsid protein	98.67	7.00E-07	Proteínas estruturais/morfogêneses
M690	302	putative thiol protease [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	Putative Cysteine protease	100	5.8e-39	Enzimas diversas
M691	549	YqaJ viral recombinase family [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	YqaJ-like viral recombinase domain	100	3.7e-48	Recombinação
M692	137	hypothetical protein [Acanthamoeba polyphaga manamvirus]	99%	2.00E-90	100.00%	sem hit	-	-	Proteína hipotética
M693	296	hypothetical protein MIMI_gp0381 [Acanthamoeba polyphaga manamvirus]	98%	0.0	100.00%	Capsid protein	99.94	7.5e-26	Proteínas estruturais/morfogêneses
M694	88	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	3.00E-55	100.00%	sem hit	-	-	Proteína hipotética
M695	120	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	3.00E-76	98.32%	sem hit	-	-	Proteína hipotética
M696	269	hypothetical protein MIMI_gp0380 [Acanthamoeba polyphaga manamvirus]	69%	4.00E-111	96.24%	Putative poly(A) polymerase catalytic subunit	99.63	3.9e-16	Transcrição
M697	817	SNF2 family helicase [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	SNF2 family helicase	100	1.6e-39	Enzimas diversas
M698	968	hypothetical protein vs_R284 [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	Ubiquitin-protein ligase E3a-like	99.1	2.00E-09	Enzimas diversas
M699	705	hypothetical protein MIMI_gp0377 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Ubiquitin-protein ligase E3a-like	93.15	0.044	Enzimas diversas
M700	485	hypothetical protein MIMI_gp0376 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M701	195	hypothetical protein MIMI_gp0375 [Acanthamoeba polyphaga manamvirus]	100%	2.00E-135	100.00%	sem hit	-	-	Proteína hipotética
M702	973	putative regulator of chromosome condensation [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Putative regulator of chromosome condensation	99.95	4.6e-25	Outras proteínas
M703	83	hypothetical protein MIMI_gp0373 [Acanthamoeba polyphaga manamvirus]	100%	1.00E-51	100.00%	sem hit	-	-	Proteína hipotética
M704	526	WD40-repeat family protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	WD40-repeat family protein	100	7.2e-29	Repetições/dominos com funções diversas
M705	298	probable ribonuclease 3 [Acanthamoeba polyphaga manamvirus]	100%	0.0	100.00%	Putative ribonuclease	100	9.2e-33	Enzimas diversas
M706	198	putative ribonuclease 3 [Acanthamoeba polyphaga manamvirus]	83%	3.00E-108	97.59%	Putative ribonuclease	99.92	1E-23	Enzimas diversas
M707	255	hypothetical protein vs_R277 [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M708	496	poxvirus poly(A) polymerase catalytic subunit-like protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	poxvirus poly(A) polymerase catalytic subunit-like protein	100	1.8e-104	Transcrição
M709	126	hypothetical protein [Acanthamoeba castellanii manamvirus]	96%	1.00E-78	100.00%	Putative Acetyltransferase	95.75	0.12	Enzimas diversas
M710	173	transcription factor S-II-related protein [Acanthamoeba polyphaga manamvirus]	100%	3.00E-123	100.00%	transcription factor S-II-related protein	99.95	3.1e-26	Transcrição
M711	895	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M712	202	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	1.00E-144	100.00%	Diamine acetyltransferase 2	98.83	2.7e-7	Enzimas diversas
M713	607	hypothetical protein vs_R273 [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	ZincRING finger domain-containing protein	100	1.2e-39	Repetições/dominos com funções diversas
M714	341	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M715	187	hypothetical protein MIMI_gp0363 [Acanthamoeba polyphaga manamvirus]	100%	1.00E-126	100.00%	sem hit	-	-	Proteína hipotética
M716	203	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	3.00E-139	100.00%	sem hit	-	-	Proteína hipotética
M717	208	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	9.00E-145	100.00%	sem hit	-	-	Proteína hipotética
M718	374	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	HITMYND zinc finger-like family protein	98.11	0.0000071	Repetições/dominos com funções diversas
M719	265	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M720	184	hypothetical protein [Acanthamoeba castellanii manamvirus]	100%	8.00E-128	100.00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no HhrpEd	Prob. %	e-value	Categoria
M721	85	hypothetical protein MIM1_gp0337 [Acanthamoeba polyphaga minivírus]	100%	2,00E-55	100.00%	sem hit	-	-	Proteína hipotética
M722	193	hypothetical protein [Acanthamoeba castellani manavírus]	100%	2,00E-133	100.00%	HNH homing endonuclease	99.02	1.9e-9	Enzimas diversas
M723	125	hypothetical protein [Acanthamoeba castellani manavírus]	100%	1,00E-76	100.00%	sem hit	-	-	Proteína hipotética
M724	995	poxvirus early transcription factor-like protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	poxvirus early transcription factor-like protein	99.83	6.2e-21	Transcrição
M725	193	hypothetical protein MIM1_R327a [Acanthamoeba polyphaga minivírus]	100%	8,00E-122	95.34%	DNA-dependent RNA polymerase complex	99.78	1.2e-19	Transcrição
M726	96	hypothetical protein MIM1_R327a [Acanthamoeba polyphaga minivírus]	100%	2,00E-54	100.00%	DNA binding domain-containing protein	94.46	0.29	Repetições/dominios com funções diversas
M727	687	poxvirus early transcription factor-like protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	sem hit	-	-	Transcrição
M728	197	putative metalloproteinase WLM1 [Acanthamoeba castellani manavírus]	100%	5,00E-139	100.00%	Ubiquitin and WLM domain-containing metalloprotease	98.69	9.3e-8	Enzimas diversas
M729	348	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M730	283	Myristylated DV envelope protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Transmembrane protein	92.97	0.98	Outras proteínas
M731	1743	DNA polymerase [Acanthamoeba polyphaga minivírus]	100%	0.0	99.83%	DNA polymerase family B	100	3.4e-112	Replicação de DNA
M732	184	hypothetical protein [Acanthamoeba castellani manavírus]	100%	2,00E-128	100.00%	sem hit	-	-	Proteína hipotética
M733	253	putative methyltransferase [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Putative methyltransferase	99.3	3.9e-10	Enzimas diversas
M734	468	putative ubiquitin-specific protease [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Putative ubiquitin-specific protease	100	4.7e-39	Enzimas diversas
M735	354	probable DNA polymerase family X [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Putative DNA polymerase family X	100	9.2e-52	Replicação de DNA
M736	230	hypothetical protein [Acanthamoeba castellani manavírus]	100%	2,00E-166	100.00%	POZ domain-containing protein	97.46	0.00066	Repetições/dominios com funções diversas
M737	148	probable glucosamine 6-phosphate N-acetyltransferase [Acanthamoeba polyphaga minivírus]	100%	4,00E-102	99.32%	Putative Acetyltransferase	99.92	5.3e-22	Enzimas diversas
M738	287	Formamidopyrimidine-DNA glycosylase [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Formamidopyrimidine-DNA glycosylase	100	2.5e-56	Reparo de DNA
M739	138	hypothetical protein [Acanthamoeba castellani manavírus]	100%	9,00E-98	100.00%	sem hit	-	-	Proteína hipotética
M740	881	ribonucleoside-diphosphate reductase large subunit [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Ribonucleoside-diphosphate reductase large subunit	100	7.3e-143	Replicação de DNA
M741	417	ribonucleotide reductase small subunit [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	ribonucleotide reductase small subunit	100	6.4e-60	Replicação de DNA
M742	456	RING domain-containing protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	RING finger domain, C3HC4	97.94	0.000014	Repetições/dominios com funções diversas
M743	71	putative ring finger protein [Acanthamoeba castellani manavírus]	98%	5,00E-41	100.00%	sem hit	-	-	Enzimas diversas
M744	172	hypothetical protein MIM1_gp0339 [Acanthamoeba polyphaga minivírus]	95%	5,00E-114	100.00%	sem hit	-	-	Proteína hipotética
M745	95	hypothetical protein MIM1_gp0338 [Acanthamoeba polyphaga minivírus]	100%	8,00E-62	100.00%	sem hit	-	-	Proteína hipotética
M746	71	hypothetical protein MIM1_gp0337 [Acanthamoeba polyphaga minivírus]	100%	2,00E-45	100.00%	sem hit	-	-	Proteína hipotética
M747	311	hypothetical protein [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M748	322	putative mRNA capping enzyme [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	putative mRNA capping enzyme	99.92	5.4e-24	Transcrição
M749	292	putative protein phosphatase 2C [Acanthamoeba polyphaga minivírus]	100%	1,00E-117	100.00%	putative protein phosphatase 2C	99.96	7.4e-27	Enzimas diversas
M750	169	hypothetical protein [Acanthamoeba castellani manavírus]	100%	2,00E-178	100.00%	sem hit	-	-	Proteína hipotética
M751	252	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M752	636	NAD-dependent DNA ligase [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	NAD-dependent DNA ligase	100	1E-103	Enzimas diversas
M753	141	hypothetical protein [Acanthamoeba castellani manavírus]	100%	3,00E-95	100.00%	sem hit	-	-	Proteína hipotética
M754	70	putative serine/threonine protein kinase [Hirudovirus strain Sangsue]	100%	3,00E-40	100.00%	sem hit	-	-	Enzimas diversas
M755	311	putative serine/threonine-protein kinase [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	putative serine/threonine-protein kinase	99.92	5.6e-24	Enzimas diversas
M756	166	uncharacterized endonuclease amino terminal domain protein [AFMV]	100%	2,00E-117	100.00%	GIY-YIG catalytic domain endonuclease	98.61	5.9e-8	Enzimas diversas
M757	70	hypothetical protein MIM1_gp0327 [Acanthamoeba polyphaga minivírus]	100%	1,00E-40	100.00%	sem hit	-	-	Proteína hipotética
M758	306	probable ribonuclease H protein [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Putative Ribonuclease H-like	99.5	1.9e-11	Enzimas diversas
M759	284	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	RNA-binding domain-containing protein	92.28	0.85	Repetições/dominios com funções diversas
M760	295	apurinic endonuclease IV [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Endonuclease IV	100	2E-34	Reparo de DNA
M761	84	hypothetical protein [Acanthamoeba polyphaga minivírus]	92%	5,00E-45	100.00%	sem hit	-	-	Proteína hipotética
M762	659	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M763	958	hypothetical protein [Acanthamoeba castellani manavírus]	99%	0.0	99.46%	sem hit	-	-	Proteína hipotética
M764	1365	hypothetical protein [Acanthamoeba castellani manavírus]	99%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M765	452	peptidase C19 subfamily protein [Acanthamoeba castellani manavírus]	100%	4,00E-132	100.00%	P-loop containing nucleoside triphosphate hydrolases superfamily protein	98.21	0.000064	Enzimas diversas
M766	191	hypothetical protein MIM1_gp0320 [Acanthamoeba polyphaga minivírus]	100%	0.0	100.00%	Cysteine proteinase	99.64	5.6e-16	Enzimas diversas
M767	577	putative low complexity protein [Acanthamoeba castellani manavírus]	100%	0.0	98.97%	sem hit	-	-	Proteína hipotética
M768	548	DEAD/SNF2-like helicase [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	DEAD/SNF2-like helicase	100	1E-50	Enzimas diversas
M769	455	ankyrin repeat-containing protein [Acanthamoeba castellani manavírus]	100%	3,00E-50	100.00%	Ankyrin repeat containing protein	100	1.1e-46	Repetições/dominios com funções diversas
M770	86	ankyrin-containing protein [Hirudovirus strain Sangsue]	100%	3,00E-150	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M771	210	Ricin-type lectin protein [Acanthamoeba castellani manavírus]	100%	2,00E-147	100.00%	Ricin-type lectin protein	97.74	0.0095	Outras proteínas
M772	209	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M773	250	hypothetical protein [Acanthamoeba castellani manavírus]	100%	6,00E-180	100.00%	Leucine rich repeat-containing protein	98.05	0.00045	Repetições/dominios com funções diversas
M774	205	hypothetical protein Ivs_L234 [Acanthamoeba polyphaga lentivírus]	100%	6,00E-154	100.00%	sem hit	-	-	Proteína hipotética
M775	356	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.97	1.5e-29	Repetições/dominios com funções diversas
M776	559	hypothetical protein [Acanthamoeba castellani manavírus]	99%	0.0	100.00%	Ankyrin repeat containing protein	99.93	1.6e-24	Repetições/dominios com funções diversas
M777	152	hypothetical protein [Acanthamoeba castellani manavírus]	100%	1,00E-107	100.00%	sem hit	-	-	Proteína hipotética
M778	292	hypothetical protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Leucine rich FNIP repeat-containing protein	99.54	7.2e-12	Repetições/dominios com funções diversas
M779	197	putative acetyltransferase [Acanthamoeba castellani manavírus]	100%	3,00E-141	100.00%	N-acetyltransferase	99.33	3.1e-12	Enzimas diversas
M780	470	ankyrin repeat-containing protein [Acanthamoeba castellani manavírus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	3.5e-44	Repetições/dominios com funções diversas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M781	126	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5,00E-88	100,00%	sem hit	-	-	Proteína hipotética
M782	129	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4,00E-88	100,00%	5 polynucleotide kinase-3' phosphatase	99,62	2,8e-13	Enzimas diversas
M783	93	putative oxoglutarate malate carrier protein [Acanthamoeba polyphaga minivirus]	89%	1,00E-47	98,78%	mitochondrial carrier protein	97,3	0,0003	Outras proteínas
M784	147	putative oxoglutarate malate carrier protein [Acanthamoeba castellanii mannavirus]	100%	8,00E-102	100,00%	mitochondrial carrier protein	99,4	7,2e-15	Outras proteínas
M785	98	hypothetical protein MIM1_gp0302 [Acanthamoeba polyphaga minivirus]	100%	1,00E-64	98,98%	sem hit	-	-	Proteína hipotética
M786	175	divergent BTB domain-protein [Acanthamoeba castellanii mannavirus]	97%	8,00E-117	100,00%	POZ domain-containing protein	99,68	9,7e-16	Repetições/dominios com funções diversas
M787	143	putative BTB-POZ domain-containing protein [Acanthamoeba polyphaga minivirus]	100%	2,00E-91	99,28%	sem hit	-	-	Repetições/dominios com funções diversas
M788	64	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-39	100,00%	sem hit	-	-	Proteína hipotética
M789	212	hypothetical protein MIM1_gp0299[Acanthamoeba polyphaga minivirus]	100%	2,00E-158	100,00%	sem hit	-	-	Proteína hipotética
M790	446	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M791	285	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	0,0	100,00%	Ankyrin repeat containing protein	100	3,4e-32	Repetições/dominios com funções diversas
M792	382	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Serine threonine-protein kinase	99,85	3,3e-21	Enzimas diversas
M793	86	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4,00E-54	100,00%	WD40 repeat-containing protein	97,57	0,0046	Repetições/dominios com funções diversas
M794	725	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M795	143	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-99	99,30%	sem hit	-	-	Proteína hipotética
M796	205	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	5,00E-142	100,00%	sem hit	-	-	Proteína hipotética
M797	191	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-132	100,00%	serine-threonine kinase	99,97	6,3e-28	Enzimas diversas
M798	343	serine-threonine kinase [Acanthamoeba castellanii mannavirus]	100%	0,0	99,71%	serine-threonine kinase	100	1,4e-54	Enzimas diversas
M799	433	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas
M800	181	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	5,00E-124	100,00%	Ankyrin repeat containing protein	99,19	4,5e-10	Repetições/dominios com funções diversas
M801	64	Dnaj-like protein [Acanthamoeba castellanii mannavirus]	-	-	-	Dnaj-like protein	-	-	Chaperona
M802	92	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M803	76	hypothetical protein [Samba virus]	100%	7,00E-55	100,00%	sem hit	-	-	Proteína hipotética
M804	1396	putative low complexity protein [Hirudovirus strain Sangsue]	100%	9,00E-44	100,00%	sem hit	-	-	Outras proteínas
M805	185	putative WD repeat-containing protein [Hirudovirus strain Sangsue]	100%	0,0	98,87%	sem hit	-	-	Repetições/dominios com funções diversas
M806	104	hypothetical protein MIM1_L263b [Acanthamoeba polyphaga minivirus]	97%	4,00E-119	97,78%	Putative WD repeat-containing protein	98,42	0,00025	Repetições/dominios com funções diversas
M807	313	hypothetical protein [Acanthamoeba castellanii mannavirus]	93%	1,00E-58	100,00%	sem hit	-	-	Proteína hipotética
M808	105	hypothetical protein [Acanthamoeba polyphaga minivirus]	99%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M809	181	hypothetical protein MIM1_L262 [Acanthamoeba polyphaga minivirus]	100%	6,00E-70	100,00%	sem hit	-	-	Proteína hipotética
M810	398	Dnaj-like protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-130	100,00%	sem hit	-	-	Outras proteínas
M811	292	putative rRNA-(His) guanlyltransferase [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	putative rRNA-(His) guanlyltransferase	99,81	3,7e-18	Chaperona
M812	172	Thymidine kinase [Acanthamoeba castellanii mannavirus]	100%	1,00E-118	100,00%	sem hit	-	-	Enzimas diversas
M813	339	hypothetical protein MIM1_gp0281[Acanthamoeba polyphaga minivirus]	99%	0,0	100,00%	YpsA-like protein	100	3,9e-38	Enzimas diversas
M814	213	hypothetical protein MIM1_gp0280[Acanthamoeba polyphaga minivirus]	100%	2,00E-148	100,00%	Dnaj-like protein	100	3,4e-57	Chaperona
M815	312	hypothetical protein MIM1_gp0279[Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Enzimas diversas
M816	917	heat shock protein 70-like protein [Acanthamoeba polyphaga minivirus]	100%	0,0	99,67%	heat shock protein 70-like protein	99,94	2,9e-26	Enzimas diversas
M817	58	hypothetical protein MIM1_gp0277[Acanthamoeba polyphaga minivirus]	100%	3,00E-33	100,00%	sem hit	-	-	Proteína hipotética
M818	176	hypothetical protein MIM1_gp0276[Acanthamoeba polyphaga minivirus]	94%	1,00E-111	100,00%	sem hit	-	-	Proteína hipotética
M819	195	hypothetical protein HIRU_S710 [Hirudovirus strain Sangsue]	93%	1,00E-127	99,45%	sem hit	-	-	Proteína hipotética
M820	268	hypothetical protein MIM1_gp0275[Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	DNA binding protein	99,8	1,8e-19	Repetições/dominios com funções diversas
M821	100	Lon protease homolog [Acanthamoeba polyphaga minivirus]	100%	1,00E-58	100,00%	sem hit	-	-	Enzimas diversas
M822	915	lon protease-like protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	lon protease-like protein	100	4,7e-53	Enzimas diversas
M823	80	hypothetical protein MIM1_gp0273[Acanthamoeba polyphaga minivirus]	100%	1,00E-51	100,00%	sem hit	-	-	Proteína hipotética
M824	445	putative transcription initiation factor IIb-like protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	putative transcription initiation factor IIb-like protein	100	2,1e-35	Transcrição
M825	370	putative uracil-DNA glycosylase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	putative uracil-DNA glycosylase	100	7,7e-80	Reparo de DNA
M826	406	hypothetical protein MIM1_gp0270[Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M827	524	DNA directed RNA polymerase subunit 2 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DNA directed RNA polymerase subunit 2	100	2,8e-70	Transcrição
M828	394	HNH endonuclease [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	HNH endonuclease	99,93	1,1e-24	Enzimas diversas
M829	666	DNA directed RNA polymerase subunit 2 [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DNA directed RNA polymerase subunit 2	100	1,8e-129	Transcrição
M830	330	uncharacterized HNH endonuclease [Acanthamoeba polyphaga minivirus]	100%	0,0	99,70%	HNH endonuclease	99,95	5E-26	Enzimas diversas
M831	109	DNA directed RNA polymerase subunit 2 [Hirudovirus strain Sangsue]	100%	1,00E-75	100,00%	DNA directed RNA polymerase subunit 2	99,94	1,6e-26	Transcrição
M832	168	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-120	100,00%	sem hit	-	-	Proteína hipotética
M833	163	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-114	100,00%	sem hit	-	-	Proteína hipotética
M834	812	collagen triple helix repeat containing protein [Acanthamoeba polyphaga lentivirus]	100%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M835	817	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M836	906	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M837	441	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M838	289	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M839	72	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	70%	5,00E-15	95,92%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M840	85	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-52	98,81%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTP	Query cover	e-value	Id. %	Best hit no Hlprep	Prob.%	e-value	Categoria
M841	205	RNA polymerase subunit 5 [Acanthamoeba castellanii mannavirus]	100%	5,00E-146	100,00%	RNA polymerase subunit 5	100	4.4e-45	Transcrição
M842	288	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M843	181	putative Zn-dependent peptidase [Acanthamoeba castellanii mannavirus]	100%	2,00E-126	100,00%	putative Zn-dependent peptidase	99.98	1.8e-29	Enzimas diversas
M844	115	putative Zn-dependent peptidase [Acanthamoeba castellanii mannavirus]	100%	7,00E-73	100,00%	putative Zn-dependent peptidase	99.53	7.1e-13	Enzimas diversas
M845	473	serine threonine protein kinase [Acanthamoeba polyphaga minivirus]	100%	1,00E-173	100,00%	serine threonine protein kinase	99.97	5.5e-30	Enzimas diversas
M846	185	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9,00E-128	100,00%	sem hit	-	-	Proteína hipotética
M847	895	Procollagen-lysine,2-oxoglutarate 5-dioxygenase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Procollagen-lysine,2-oxoglutarate 5-dioxygenase	100	1.5e-72	Enzimas diversas
M848	290	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	99.94	5.7e-25	Repetições/dominios com funções diversas
M849	379	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	5.7e-34	Repetições/dominios com funções diversas
M850	212	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4,00E-143	100,00%	Ankyrin repeat containing protein	99.91	2.3e-23	Repetições/dominios com funções diversas
M851	170	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-119	100,00%	sem hit	-	-	Proteína hipotética
M852	691	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	99,42%	Ankyrin repeat containing protein	100	1.3e-51	Repetições/dominios com funções diversas
M853	531	N-terminal divergent BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	WD40 repeat-containing protein	99.95	1.5e-24	Repetições/dominios com funções diversas
M854	435	putative BTB POZ domain-containing protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	WD40 repeat-containing protein	100	8.1e-30	Repetições/dominios com funções diversas
M855	92	putative BTB POZ domain-containing protein [Acanthamoeba polyphaga minivirus]	97%	9,00E-53	100,00%	putative BTB POZ domain-containing protein	99.3	1.7e-10	Repetições/dominios com funções diversas
M856	156	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-109	100,00%	sem hit	-	-	Proteína hipotética
M857	308	Protein lysine methyltransferase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Protein lysine methyltransferase	99.78	8.5e-17	Enzimas diversas
M858	828	Topoisomerase I [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Topoisomerase I	100	3.3e-107	Enzimas diversas
M859	340	hypothetical protein MIM1_gp0242[Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	sem hit	-	-	Proteína hipotética
M860	663	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	0,0	100,00%	Ankyrin repeat containing protein	100	4.8e-34	Repetições/dominios com funções diversas
M861	258	hypothetical protein MIM1_gp0240[Acanthamoeba polyphaga minivirus]	96%	1,00E-179	98,80%	sem hit	-	-	Proteína hipotética
M862	732	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	ADP-ribosylation superfamily protein	99.75	4.9e-17	Outras proteínas
M863	226	putative ankyrin repeat protein [Acanthamoeba polyphaga minivirus]	100%	1,00E-159	100,00%	Ankyrin repeat containing protein	100	1.2e-33	Repetições/dominios com funções diversas
M864	183	Rab family GTPase [Acanthamoeba castellanii mannavirus]	96%	4,00E-125	100,00%	Rab family GTPase	99.96	1.3e-26	Proteínas estruturas/morfogêneses
M865	174	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-122	100,00%	sem hit	-	-	Proteína hipotética
M866	276	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	99,64%	sem hit	-	-	Proteína hipotética
M867	262	hypothetical protein vs. L1_64 [Acanthamoeba polyphaga lentillevirus]	100%	0,0	100,00%	Transmembrane protein	99.91	5.8e-22	Outras proteínas
M868	383	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Putative zincin-like metalloprotease	93.54	0.13	Enzimas diversas
M869	387	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	DNA-directed RNA polymerase subunit	99.88	3.8e-22	Transcrição
M870	192	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-138	100,00%	DNA-directed RNA polymerase subunit	99.36	3.4e-12	Transcrição
M871	960	primase/helicase [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	primase/helicase	99.96	1.1e-28	Replicação de DNA
M872	531	Protein kinase family protein [Acanthamoeba castellanii mannavirus]	100%	0,0	100,00%	Protein kinase family protein	100	1.1e-34	Enzimas diversas
M873	90	hypothetical protein MIM1_gp0227[Acanthamoeba polyphaga minivirus]	71%	1,00E-37	100,00%	sem hit	-	-	Proteína hipotética
M874	103	hypothetical protein MIM1_L204b [Acanthamoeba polyphaga minivirus]	100%	2,00E-67	99,03%	sem hit	-	-	Proteína hipotética
M875	191	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-124	100,00%	sem hit	-	-	Proteína hipotética
M876	180	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-126	100,00%	sem hit	-	-	Proteína hipotética
M877	92	hypothetical protein HIRU_S765 [Hirudovirus strain Sangsue]	100%	1,00E-46	100,00%	sem hit	-	-	Proteína hipotética
M878	158	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1,00E-106	100,00%	sem hit	-	-	Proteína hipotética
M879	234	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2,00E-167	100,00%	sem hit	-	-	Proteína hipotética
M880	115	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3,00E-75	100,00%	sem hit	-	-	Proteína hipotética
M881	181	putative cytidine deaminase [Acanthamoeba castellanii mannavirus]	100%	4,00E-126	100,00%	Cytidine deaminase-like protein	99.74	4.4e-17	Outras proteínas
M882	1360	collagen triple helix repeat-containing protein [Acanthamoeba polyphaga lentillevirus]	100%	0,0	100,00%	collagen triple helix repeat-containing protein	100	4.1e-43	Repetições/dominios com funções diversas (colágeno)
M883	77	collagen triple helix repeat-containing protein [Acanthamoeba polyphaga minivirus]	100%	1,00E-46	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M884	719	collagen triple helix repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	0,0	100,00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M885	110	hypothetical protein [Niameyex virus]	83%	3,00E-54	98,90%	sem hit	-	-	Proteína hipotética
M886	151	hypothetical protein MIM1_gp0224[Acanthamoeba polyphaga minivirus]	100%	4,00E-106	100,00%	sem hit	-	-	Proteína hipotética
M887	285	hypothetical protein HIRU_S765 [Hirudovirus strain Sangsue]	100%	0,0	99,65%	sem hit	-	-	Proteína hipotética
M888	152	V12 [Sputnik virophage]	100%	1,00E-99	100,00%	sem hit	-	-	Proteína hipotética
M889	165	hypothetical protein Sputnik2_R11 [Sputnik virophage 2]	100%	4,00E-110	99,39%	sem hit	-	-	Proteína hipotética
M890	226	V10 [Sputnik virophage]	100%	2,00E-163	100,00%	transposases protein-DNA complex, tyrosine recombinase	99.89	4.4e-21	Proteína hipotética (vírofago)
M891	175	V9 [Sputnik virophage]	100%	1,00E-124	100,00%	Cysteine proteinases superfamily	99.82	2.2e-19	Recombinação (vírofago)
M892	184	V8 [Sputnik virophage]	100%	8,00E-126	100,00%	sem hit	-	-	Morfogênese (vírofago)
M893	236	V7 [Sputnik virophage]	100%	5,00E-162	100,00%	sem hit	-	-	Morfogênese (vírofago)
M894	310	V6 [Sputnik virophage]	100%	0,0	100,00%	COLLAGEN-LIKE PROTEIN	98.88	0.0000017	Repetições/dominios com funções diversas (vírofago)
M895	119	V5 [Sputnik virophage]	100%	3,00E-75	100,00%	Collagen alpha-2(I) chain	97.65	0.0033	Repetições/dominios com funções diversas (vírofago)
M896	139	V4 [Sputnik virophage]	100%	1,00E-99	100,00%	sem hit	-	-	Proteína hipotética (vírofago)
M897	245	V3 [Sputnik virophage]	100%	1,00E-178	100,00%	C2H2-type zinc finger	97.86	0.000017	Ligação ao DNA (vírofago)
M898	114	V2 [Sputnik virophage]	100%	9,00E-75	100,00%	FtsK-HerA superfamily, P-loop ATPase	99.53	2.9e-12	Morfogênese (vírofago)
M899	144	V1 [Sputnik virophage]	100%	7,00E-94	100,00%	sem hit	-	-	Proteína hipotética (vírofago)
M900	442	V21 [Sputnik virophage]	100%	0,0	100,00%	C4-type zinc ribbon domain Transferrase	87.68	0.54	Ligação ao DNA (vírofago)

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M901	345	V20 [Sputnik viroplage]	100%	0.0	100.00%	capsid protein V20	100	6.8e-139	Morfogênese (virofago)
M902	378	minor virion protein [Sputnik viroplage 2]	100%	0.0	100.00%	Minor virion protein	100	8.4e-112	Morfogênese (virofago)
M903	189	putative IS3 family transposase A protein [Sputnik viroplage 2]	99%	4.00E-129	100.00%	Transcriptional regulator (TetR-type HTH domain)	97.24	0.0016	Ligação ao DNA (virofago)
M904	130	V16 [Sputnik viroplage]	100%	1.00E-88	100.00%	sem hit	-	-	Proteína hipotética (virofago)
M905	109	V15 [Sputnik viroplage]	100%	9.00E-71	100.00%	sem hit	-	-	Otras proteínas (virofago)
M906	114	V14 [Sputnik viroplage]	100%	1.00E-71	100.00%	Zinc motif and AAA domain	96.56	0.053	Ligação ao DNA (virofago)
M907	779	V13 [Sputnik viroplage]	100%	0.0	100.00%	Primase, helicase	99.95	8.3e-27	Replicação de DNA (virofago)
M908	71	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	65%	1.00E-11	91.11%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M909	68	collagen-like protein 6 [Niameyer virus]	95%	3.00E-07	48.68%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M910	396	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	94%	0.0	100.00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M911	102	probable glutaredoxin [Acanthamoeba polyplaga minivirus]	100%	3.00E-69	100.00%	Putative glutaredoxin	99.33	4.9e-10	Enzimas diversas
M912	336	DNA topoisomerase 1b [Acanthamoeba polyplaga minivirus]	100%	0.0	100.00%	DNA topoisomerase 1b	100	5E-80	Enzimas diversas
M913	601	putative glycosyl-transferase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative glycosyl-transferase	99.97	7.3e-27	Enzimas diversas
M914	255	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-30	100.00%	DEAD/DEAH box helicase domain	99.88	1.3e-21	Enzimas diversas
M915	54	putative membrane protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Otras proteínas
M916	78	hypothetical protein MIM1_gp0212[Acanthamoeba polyplaga minivirus]	100%	1.00E-47	100.00%	sem hit	-	-	Proteína hipotética
M917	276	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M918	359	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	ankyrin repeat containing protein	99.97	3.3e-29	Repetições/dominios com funções diversas
M919	112	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	4.00E-69	100.00%	sem hit	-	-	Proteína hipotética
M920	509	Flavin-containing amine oxidoreductase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Flavin-containing amine oxidoreductase	100	4.4e-40	Enzimas diversas
M921	246	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M922	482	putative transposase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative transposase DNA-binding domain	99.97	1.5e-29	Enzimas diversas
M923	320	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M924	194	hypothetical protein [Acanthamoeba castellanii mannavirus]	76%	3.00E-91	94.56%	sem hit	-	-	Proteína hipotética
M925	65	hypothetical protein MIM1_L183 [Acanthamoeba polyplaga minivirus]	100%	1.00E-34	100.00%	sem hit	-	-	Proteína hipotética
M926	335	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	99.40%	sem hit	-	-	Proteína hipotética
M927	389	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M928	150	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-103	100.00%	sem hit	-	-	Proteína hipotética
M929	180	hypothetical protein HIRU_S787 [Hirudovirus strain Sangsue]	100%	7.00E-120	100.00%	ankyrin repeat containing protein	92.44	0.47	Repetições/dominios com funções diversas
M930	367	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	ankyrin repeat containing protein	93.45	0.29	Repetições/dominios com funções diversas
M931	379	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	ankyrin repeat containing protein	91.56	0.25	Repetições/dominios com funções diversas
M932	378	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	ankyrin repeat containing protein	93.65	0.091	Repetições/dominios com funções diversas
M933	242	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-167	100.00%	sem hit	-	-	Proteína hipotética
M934	149	hypothetical protein MIM1_L174 [Acanthamoeba polyplaga minivirus]	100%	4.00E-67	81.17%	ankyrin repeat containing protein	89.66	0.57	Repetições/dominios com funções diversas
M935	103	hypothetical protein HIRU_S792 [Hirudovirus strain Sangsue]	94%	3.00E-64	98.97%	sem hit	-	-	Proteína hipotética
M936	169	putative ATP-dependent RNA helicase [Acanthamoeba polyplaga minivirus]	100%	5.00E-114	97.02%	putative ATP-dependent RNA helicase	99.43	1.7e-11	Enzimas diversas
M937	224	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-159	100.00%	sem hit	-	-	Proteína hipotética
M938	100	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-61	100.00%	ankyrin repeat containing protein	98.81	2.1e-7	Repetições/dominios com funções diversas
M939	151	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-105	100.00%	sem hit	-	-	Proteína hipotética
M940	437	FNIP repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Leucine rich and FNIP repeat-containing protein	99.79	1.6e-15	Repetições/dominios com funções diversas
M941	154	hypothetical protein [Acanthamoeba castellanii mannavirus]	91%	4.00E-99	100.00%	sem hit	-	-	Proteína hipotética
M942	307	hypothetical protein [Acanthamoeba castellanii mannavirus]	83%	1.00E-180	98.10%	Leucine rich and FNIP repeat-containing protein	97.96	0.00023	Repetições/dominios com funções diversas
M943	251	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Leucine rich and FNIP repeat-containing protein	95.25	0.031	Repetições/dominios com funções diversas
M944	271	hypothetical protein vs L124 [Acanthamoeba polyplaga lentivirus]	100%	0.0	100.00%	Leucine rich and FNIP repeat-containing protein	98.28	0.0000067	Repetições/dominios com funções diversas
M945	509	Cysteinylyl-RNA synthetase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Cysteinylyl-RNA synthetase	100	1.5e-53	Tradução
M946	130	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	4.00E-93	100.00%	sem hit	-	-	Proteína hipotética
M947	521	putative FNIP repeat-containing protein [Acanthamoeba polyplaga minivirus]	100%	0.0	100.00%	Leucine rich and FNIP repeat-containing protein	99.8	2.3e-15	Repetições/dominios com funções diversas
M948	59	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	4.00E-32	100.00%	sem hit	-	-	Proteína hipotética
M949	319	hypothetical protein [Acanthamoeba castellanii minivirus]	98%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M950	289	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M951	88	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	2.00E-54	100.00%	sem hit	-	-	Proteína hipotética
M952	194	putative transposase [Acanthamoeba polyplaga minivirus]	87%	1.00E-113	100.00%	Putative transposase DNA-binding domain	99.85	3.9e-20	Enzimas diversas
M953	105	putative transposase [Acanthamoeba polyplaga lentivirus]	100%	7.00E-63	100.00%	sem hit	-	-	Enzimas diversas
M954	138	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-97	100.00%	sem hit	-	-	Proteína hipotética
M955	370	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M956	522	hypothetical protein [Acanthamoeba castellanii mannavirus]	96%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M957	253	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Alkylated DNA repair protein alkB homologue 2	99.92	2.6e-23	Proteína hipotética
M958	142	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-98	100.00%	sem hit	-	-	Proteína hipotética
M959	467	WD40 superfamily protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	WD40 repeat-containing protein	99.95	1.9e-24	Repetições/dominios com funções diversas
M960	152	hypothetical protein [Acanthamoeba castellanii minivirus]	100%	2.00E-105	100.00%	Uncharacterized protein DUF305	99.28	9.5e-11	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrp2	Prob.%	e-value	Categoria
M961	572	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0.0	100.00%	Putative A ₁ P-binding enzyme	99.92	7.4e-24	Enzimas diversas
M962	206	N-acetyltransferase [Acanthamoeba castellanii mannavirus]	100%	7.00E-149	100.00%	N-acetyltransferase	100	1.2e-42	Enzimas diversas
M963	266	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative ribonuclease	97.74	0.000087	Enzimas diversas
M964	173	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-119	100.00%	sem hit	-	-	Proteína hipotética
M965	229	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-164	100.00%	Ankyrin repeat containing protein	99.95	4E-26	Repetições/dominos com funções diversas
M966	286	hypothetical protein [Acanthamoeba castellanii mannavirus]	97%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M967	346	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M968	274	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	7.00E-43	96.00%	sem hit	-	-	Proteína hipotética
M969	84	hypothetical protein [Acanthamoeba castellanii mannavirus]	89%	8.00E-31	100.00%	sem hit	-	-	Proteína hipotética
M970	58	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-31	100.00%	sem hit	-	-	Proteína hipotética
M971	287	putative exoV-like protein [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	Polysaccharide pyruvyl transferase	99.91	5.4e-23	Enzimas diversas
M972	490	putative WbbJ Acetyltransferase [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	UDP N-Acetylglucosamine Acetyltransferase domain-containing protein	100	8.1e-34	Enzimas diversas
M973	323	putative dTDP-D-glucose 4,6-dehydratase [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	putative dTDP-D-glucose 4,6-dehydratase	100	1.6e-40	Enzimas diversas
M974	304	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Glycosyltransferase	99.84	1.4e-18	Enzimas diversas
M975	248	uncharacterized glycosyltransferase [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	Putative glycosyltransferase	99.95	2.1e-25	Enzimas diversas
M976	883	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Putative glycosyltransferase	99.63	2.5e-15	Enzimas diversas
M977	702	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	beta-Glycosyltransferase	99.68	5.7e-15	Enzimas diversas
M978	352	Aminotransferase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Aminotransferase	100	2.1e-31	Enzimas diversas
M979	702	choline dehydrogenase-like protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-79	96.72%	GMC oxidoreductase domain-containing protein	100	4.3e-59	Enzimas diversas
M980	122	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M981	93	hypothetical protein [Acanthamoeba castellanii mannavirus]	95%	2.00E-49	89.89%	sem hit	-	-	Proteína hipotética
M982	127	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-84	100.00%	sem hit	-	-	Proteína hipotética
M983	221	Methyltransferase family protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-161	100.00%	Methyltransferase family protein	99.67	1.3e-14	Enzimas diversas
M984	178	hypothetical protein vs_R89 [Acanthamoeba polyphaga lentivirus]	100%	1.00E-124	100.00%	sem hit	-	-	Proteína hipotética
M985	421	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M986	70	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-43	100.00%	sem hit	-	-	Proteína hipotética
M987	563	Choline dehydrogenase [Acanthamoeba castellanii mannavirus]	100%	0.0	99.47%	GMC oxidoreductase domain-containing protein	100	2.6e-48	Enzimas diversas
M988	269	putative F-box protein [Acanthamoeba polyphaga minnavirus]	100%	0.0	99.26%	putative F-box domain-containing protein	98.13	0.000018	Repetições/dominos com funções diversas
M989	255	putative F-box protein [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	putative F-box domain-containing protein	98.11	0.00002	Repetições/dominos com funções diversas
M990	196	BRO-M family protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-140	100.00%	sem hit	-	-	repetições/dominos com funções diversas
M991	346	Tyrosyl-RNA synthetase [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Tyrosyl-RNA synthetase	100	1.5e-51	Tradução
M992	83	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-56	98.80%	zinc finger protein YacG	90.9	0.35	Repetições/dominos com funções diversas
M993	627	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	3.8e-51	Repetições/dominos com funções diversas
M994	284	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	98%	0.0	100.00%	Ankyrin repeat containing protein	100	4.7e-31	Repetições/dominos com funções diversas
M995	339	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	99.71%	Ankyrin repeat containing protein	100	1.2e-36	Repetições/dominos com funções diversas
M996	75	hypothetical protein HIRU_S850 [Hirudovirus strain Sanguis]	100%	6.00E-48	100.00%	sem hit	-	-	Proteína hipotética
M997	165	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-116	100.00%	sem hit	-	-	Proteína hipotética
M998	356	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-65	100.00%	sem hit	99.84	4.3e-18	Enzimas diversas
M999	97	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	MORN repeat-containing protein	98.82	7.8e-9	Repetições/dominos com funções diversas
M1000	535	Helicase, SNF2 family [Acanthamoeba castellanii mannavirus]	100%	6.00E-43	100.00%	SNF2 family helicase	99.97	1.3e-28	Enzimas diversas
M1001	1262	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-84	100.00%	Rubredoxin-like fold containing protein	85.06	0.57	Outras proteínas
M1002	67	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-128	100.00%	sem hit	-	-	Proteína hipotética
M1003	124	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	2E-42	Repetições/dominos com funções diversas
M1004	183	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	MORN repeat-containing protein	98.9	2.00E-08	Repetições/dominos com funções diversas
M1005	458	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	6.00E-104	100.00%	sem hit	-	-	Proteína hipotética
M1006	374	hypothetical protein [Acanthamoeba castellanii mannavirus]	90%	1.00E-84	100.00%	Ankyrin repeat containing protein	99.48	2.7e-12	Repetições/dominos com funções diversas
M1007	166	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-170	100.00%	Ankyrin repeat containing protein	99.9	4.9e-22	Repetições/dominos com funções diversas
M1008	128	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-170	97.25%	Proliferating cell nuclear antigen	99.97	4E-29	Replicação de DNA
M1009	238	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	POZ domain containing protein	94.94	0.24	Repetições/dominos com funções diversas
M1010	249	BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	dihal1 domain phosphoesterase	99.77	6.5e-19	Enzimas diversas
M1011	498	BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	83%	2.00E-71	100.00%	dihal1 domain phosphoesterase	99.26	2.7e-11	Enzimas diversas
M1012	312	Phosphoesterase, dihal1 [Acanthamoeba castellanii mannavirus]	99%	6.00E-66	99.00%	BTB POZ domain-containing protein	99.75	1.8e-17	Repetições/dominos com funções diversas
M1013	132	Phosphoesterase, dihal1 [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1014	167	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Mut-Like protein	96.52	0.0074	Reparo de DNA
M1015	107	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	2E-33	Repetições/dominos com funções diversas
M1016	344	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	2.5e-35	Repetições/dominos com funções diversas
M1017	599	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	WD40 repeat-containing protein	97.6	0.033	Repetições/dominos com funções diversas
M1018	272	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	DEAD/DEAH box helicase domain	99.93	3.2e-25	Enzimas diversas
M1019	503	BTB domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%				
M1020	279	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%				

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhrped	Prob. %	e-value	Categoria
M1021	193	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-133	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M1022	718	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	1.3e-53	Repetições/dominios com funções diversas
M1023	82	putative BTB-POZ domain-containing protein [Acanthamoeba polyphaga mimivirus]	87%	3.00E-41	100.00%	BTB-POZ domain-containing protein	98.94	5.6e-9	Repetições/dominios com funções diversas
M1024	412	putative BTB-POZ domain-containing protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	WD40 repeat-containing protein	97.6	0.033	Repetições/dominios com funções diversas
M1025	94	putative transposase [Acanthamoeba castellanii mannavirus]	100%	9.00E-62	100.00%	Putative transposase DNA-binding domain	99.34	1.2e-12	Enzimas diversas
M1026	879	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-124	100.00%	Ankyrin repeat containing protein	100	1.9e-65	Repetições/dominios com funções diversas
M1027	171	hypothetical protein [Acanthamoeba castellanii mannavirus]	99%	0.0	94.24%	sem hit	-	-	Proteína hipotética
M1028	890	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	1.9e-42	Repetições/dominios com funções diversas
M1029	327	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-126	100.00%	Ankyrin repeat containing protein	100	1.4e-36	Repetições/dominios com funções diversas
M1030	175	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1031	808	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	1.7e-62	Repetições/dominios com funções diversas
M1032	78	resolvase [Samba virus]	100%	3.00E-47	100.00%	Putative resolvase	95.26	0.18	Recombinação
M1033	59	putative resolvase [Acanthamoeba polyphaga lentivirus]	100%	1.00E-34	100.00%	Putative DNA-binding domain	97.59	0.000011	Recombinação
M1034	76	putative transposase [Acanthamoeba polyphaga lentivirus]	98%	4.00E-46	100.00%	sem hit	-	-	Enzimas diversas
M1035	231	putative transposase [Hirudovirus strain Saugseu]	88%	3.00E-139	100.00%	sem hit	-	-	Enzimas diversas
M1036	77	putative transposase [Nimivirus Bombay]	90%	3.00E-42	100.00%	sem hit	-	-	Enzimas diversas
M1037	59	putative transposase [Nimivirus Bombay]	100%	4.00E-35	98.31%	sem hit	-	-	Enzimas diversas
M1038	80	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	1.00E-44	97.47%	sem hit	-	-	Repetições/dominios com funções diversas
M1039	539	putative BTB-POZ domain-containing protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	86.58%	WD40 repeat-containing protein	99.16	2.00E-07	Repetições/dominios com funções diversas
M1040	133	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-92	100.00%	sem hit	-	-	Proteína hipotética
M1041	210	hypothetical protein [Acanthamoeba castellanii mannavirus]	86%	1.00E-130	100.00%	sem hit	-	-	Proteína hipotética
M1042	200	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-149	100.00%	Ankyrin repeat containing protein	99.95	1.1e-26	Repetições/dominios com funções diversas
M1043	190	putative resolvase [Acanthamoeba castellanii mannavirus]	100%	4.00E-135	99.47%	Putative resolvase	99.81	4.8e-18	Recombinação
M1044	275	putative transposase [Acanthamoeba castellanii mannavirus]	96%	0.0	99.23%	sem hit	-	-	Enzimas diversas
M1045	202	putative transposase [Acanthamoeba castellanii mannavirus]	95%	4.00E-135	98.96%	Putative transposase DNA-binding domain	99.76	2.2e-18	Enzimas diversas
M1046	99	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-63	97.98%	sem hit	-	-	Proteína hipotética
M1047	103	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6.00E-68	100.00%	sem hit	-	-	Proteína hipotética
M1048	87	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	9.00E-55	100.00%	sem hit	-	-	Proteína hipotética
M1049	242	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-176	100.00%	sem hit	-	-	Proteína hipotética
M1050	339	hypothetical protein Ivs_L46 [Acanthamoeba polyphaga lentivirus]	99%	0.0	99.41%	Serine/threonine protein kinase	99.76	4.6e-17	Enzimas diversas
M1051	112	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-71	100.00%	sem hit	-	-	Proteína hipotética
M1052	382	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	99%	0.0	100.00%	Ankyrin repeat containing protein	100	9.1e-39	Repetições/dominios com funções diversas
M1053	94	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-58	100.00%	sem hit	-	-	Proteína hipotética
M1054	53	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-26	100.00%	sem hit	-	-	Proteína hipotética
M1055	308	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1056	238	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-174	100.00%	sem hit	-	-	Proteína hipotética
M1057	227	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-159	100.00%	sem hit	-	-	Proteína hipotética
M1058	132	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-90	100.00%	Poly A polymerase head domain-like	96.44	0.0035	Transcrição
M1059	384	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	1.6e-38	Repetições/dominios com funções diversas
M1060	894	Collagen triple helix repeat containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Repetições/dominios com funções diversas (colágeno)
M1061	179	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-127	100.00%	sem hit	-	-	Proteína hipotética
M1062	195	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-139	100.00%	sem hit	-	-	Proteína hipotética
M1063	243	ankyrin repeat-containing protein [Hirudovirus strain Saugseu]	100%	8.00E-174	100.00%	Ankyrin repeat containing protein	100	3E-33	Repetições/dominios com funções diversas
M1064	498	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	WD40 repeat-containing protein	99.95	2.5e-25	Repetições/dominios com funções diversas
M1065	242	Putative ankyrin repeat protein L66 [Acanthamoeba polyphaga mimivirus]	100%	7.00E-168	100.00%	Ankyrin repeat containing protein	100	6.7e-33	Repetições/dominios com funções diversas
M1066	337	ankyrin-containing protein [Acanthamoeba polyphaga lentivirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	100	2.3e-40	Repetições/dominios com funções diversas
M1067	355	uncharacterized virion-associated membrane protein [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	sem hit	-	-	Proteínas estruturais/morfogêneses
M1068	565	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	oxidoreductase	100	2.9e-82	Enzimas diversas
M1069	580	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	8.00E-46	100.00%	Ankyrin repeat containing protein	100	2.6e-48	Repetições/dominios com funções diversas
M1070	79	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	93.62%	Ankyrin repeat containing protein	99.18	2.1e-10	Repetições/dominios com funções diversas
M1071	283	ankyrin-containing protein [Acanthamoeba polyphaga mimivirus]	100%	1.00E-32	100.00%	Ankyrin repeat containing protein	100	7.7e-41	Repetições/dominios com funções diversas
M1072	62	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-103	99.32%	sem hit	-	-	Proteína hipotética
M1073	147	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	99.73%	sem hit	-	-	Proteína hipotética
M1074	367	WD-repeat family protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-83	100.00%	WD40 repeat-containing protein	100	7.7e-38	Repetições/dominios com funções diversas
M1075	132	WD-repeat family protein [Acanthamoeba castellanii mannavirus]	100%	0.0	99.84%	sem hit	-	-	Repetições/dominios com funções diversas
M1076	632	hypothetical protein M1M_L60 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	Leucine rich and FNIP repeat containing protein	100	1.6e-33	Repetições/dominios com funções diversas
M1077	297	putative ankyrin repeat protein [Acanthamoeba polyphaga mimivirus]	100%	3.00E-128	100.00%	Ankyrin repeat containing protein	100	6.1e-37	Repetições/dominios com funções diversas
M1078	182	hypothetical protein [Acanthamoeba castellanii mannavirus]	96%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1079	315	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1080	272	ankyrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat containing protein	99.97	1E-29	Repetições/dominios com funções diversas

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hhpred	Prob. %	e-value	Categoria
M1081	513	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	WD40 repeat-containing protein	99.97	4.7e-27	Repetições/dominos com funções diversas
M1082	182	hypothetical protein [Acanthamoeba polyphaga minnavirus]	100%	6.00E-54	100.00%	sem hit	-	-	Proteína hipotética
M1083	100	hypothetical protein HIRU_S925 [Hirudovirus strain Sangsue]	100%	2.00E-85	100.00%	sem hit	-	-	Proteína hipotética
M1084	127	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-113	100.00%	Proliferating cell nuclear antigen	99.89	1.5e-21	Replicação de DNA
M1085	159	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-77	100.00%	sem hit	-	-	Proteína hipotética
M1086	113	hypothetical protein [Acanthamoeba castellanii mannavirus]	93%	2.00E-102	98.67%	sem hit	-	-	Proteína hipotética
M1087	161	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-87	97.67%	Kelch repeat and BTB domain-containing protein	99.3	1.2e-10	Repetições/dominos com funções diversas
M1088	130	hypothetical protein [Acanthamoeba polyphaga minnavirus]	99%	1.00E-87	97.67%	DNA-directed RNA polymerase	95.84	0.02	Transcrição
M1089	101	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M1090	175	hypothetical protein NDM1_gp0058 [Acanthamoeba polyphaga minnavirus]	100%	4.00E-119	96.57%	DEAD/DEAH box helicase domain	99.58	2E-14	Enzimas diversas
M1091	159	hypothetical protein NDM1_gp0057 [Acanthamoeba polyphaga minnavirus]	99%	3.00E-114	100.00%	sem hit	-	-	Proteína hipotética
M1092	227	putative ankryrin repeat protein [Acanthamoeba polyphaga minnavirus]	100%	8.00E-163	100.00%	Ankyrin repeat-containing protein	99.94	6.9e-24	Repetições/dominos com funções diversas
M1093	215	hypothetical protein NDM1_gp0055 [Acanthamoeba polyphaga minnavirus]	100%	2.00E-153	100.00%	sem hit	-	-	Proteína hipotética
M1094	212	hypothetical protein NDM1_gp0054 [Acanthamoeba polyphaga minnavirus]	100%	4.00E-152	100.00%	sem hit	-	-	Proteína hipotética
M1095	31	hypothetical protein NDM1_gp0053 [Acanthamoeba polyphaga minnavirus]	100%	1.00E-28	100.00%	sem hit	-	-	Proteína hipotética
M1096	357	putative ankryrin repeat protein [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	Ankyrin repeat-containing protein	100	6E-38	Repetições/dominos com funções diversas
M1097	478	hypothetical protein NDM1_gp0051 [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	Probable ATP-dependent RNA helicase	98.12	0.0000039	Enzimas diversas
M1098	58	hypothetical protein NDM1_gp0050 [Acanthamoeba polyphaga minnavirus]	100%	1.00E-33	100.00%	Ankyrin repeat-containing protein	95.02	0.061	Repetições/dominos com funções diversas
M1099	91	hypothetical protein NDM1_gp0049 [Acanthamoeba polyphaga minnavirus]	100%	5.00E-56	100.00%	sem hit	-	-	Proteína hipotética
M1100	101	hypothetical protein NDM1_gp0048 [Acanthamoeba polyphaga minnavirus]	100%	3.00E-62	100.00%	sem hit	-	-	Proteína hipotética
M1101	174	hypothetical protein NDM1_gp0047 [Acanthamoeba polyphaga minnavirus]	100%	7.00E-126	100.00%	sem hit	-	-	Proteína hipotética
M1102	114	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	3.00E-73	99.12%	sem hit	-	-	Proteína hipotética
M1103	97	putative ankryrin repeat protein [Acanthamoeba polyphaga minnavirus]	100%	2.00E-63	100.00%	Ankyrin repeat-containing protein	99.66	3.2e-14	Repetições/dominos com funções diversas
M1104	142	putative ankryrin repeat protein [Acanthamoeba polyphaga minnavirus]	87%	1.00E-82	100.00%	Ankyrin repeat-containing protein	99.86	3.6e-20	Repetições/dominos com funções diversas
M1105	484	putative Kilia-N domain-containing protein [Acanthamoeba polyphaga minnavirus]	100%	0.0	99.38%	putative Kilia-N domain-containing protein	98.42	0.0000029	Repetições/dominos com funções diversas
M1106	117	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	6.00E-76	99.14%	Ankyrin repeat-containing protein	99.17	6.8e-10	Repetições/dominos com funções diversas
M1107	187	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	1.00E-130	99.47%	Ankyrin repeat-containing protein	100	1.7e-31	Repetições/dominos com funções diversas
M1108	481	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	WD40 repeat-containing protein	100	4E-30	Repetições/dominos com funções diversas
M1109	84	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-54	100.00%	sem hit	-	-	Proteína hipotética
M1110	340	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	96.03%	Ankyrin repeat-containing protein	99.87	1.1e-20	Repetições/dominos com funções diversas
M1111	73	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	5.00E-43	100.00%	sem hit	-	-	Proteína hipotética
M1112	230	Kilia N-domain-containing protein [Acanthamoeba castellanii mannavirus]	99%	5.00E-150	100.00%	sem hit	-	-	Repetições/dominos com funções diversas
M1113	317	Kilia N-domain-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	putative Kilia N-domain-containing protein	97.71	0.00016	Repetições/dominos com funções diversas
M1114	108	hypothetical protein NDM1_gp0032 [Acanthamoeba polyphaga minnavirus]	100%	3.00E-75	100.00%	DNA-directed RNA polymerase	95.56	0.034	Transcrição
M1115	109	hypothetical protein [Acanthamoeba castellanii mannavirus]	87%	1.00E-60	100.00%	sem hit	-	-	Proteína hipotética
M1116	104	hypothetical protein NDM1_gp0031 [Acanthamoeba polyphaga minnavirus]	100%	2.00E-67	100.00%	sem hit	-	-	Proteína hipotética
M1117	146	hypothetical protein NDM1_gp0030 [Acanthamoeba polyphaga minnavirus]	100%	5.00E-101	100.00%	sem hit	-	-	Proteína hipotética
M1118	104	hypothetical protein NDM1_gp0029 [Acanthamoeba polyphaga minnavirus]	100%	3.00E-71	100.00%	sem hit	-	-	Proteína hipotética
M1119	326	ankryrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Ankyrin repeat-containing protein	100	2.5e-47	Repetições/dominos com funções diversas
M1120	102	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-69	100.00%	sem hit	-	-	Proteína hipotética
M1121	63	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M1122	236	ankryrin repeat-containing protein [Acanthamoeba castellanii mannavirus]	100%	2.00E-168	100.00%	Ankyrin repeat-containing protein	100	3.8e-32	Repetições/dominos com funções diversas
M1123	122	putative ankryrin repeat protein [Acanthamoeba polyphaga minnavirus]	100%	3.00E-79	100.00%	Ankyrin repeat-containing protein	99.84	8.7e-18	Repetições/dominos com funções diversas
M1124	288	putative sel1-like repeat-containing protein [Acanthamoeba polyphaga minnavirus]	97%	0.0	100.00%	putative sel1-like repeat-containing protein	99.82	7.5e-17	Repetições/dominos com funções diversas
M1125	263	sel1-like repeat-containing protein [Banda megavirus]	97%	0.0	100.00%	sem hit	-	-	Repetições/dominos com funções diversas
M1126	61	No significant similarity found	-	-	-	sem hit	-	-	Proteína hipotética
M1127	112	hypothetical protein NDM1_gp0022 [Acanthamoeba polyphaga minnavirus]	100%	4.00E-74	100.00%	sem hit	-	-	Proteína hipotética
M1128	584	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	0.0	100.00%	Tetratricopeptide repeat-containing protein	99.9	5.00E-20	Repetições/dominos com funções diversas
M1129	72	hypothetical protein [Acanthamoeba castellanii mannavirus]	98%	3.00E-38	100.00%	sem hit	-	-	Proteína hipotética
M1130	181	hypothetical protein NDM1_gp0020 [Acanthamoeba polyphaga minnavirus]	100%	2.00E-124	100.00%	sem hit	-	-	Proteína hipotética
M1131	269	hypothetical protein NDM1_gp0019 [Acanthamoeba polyphaga minnavirus]	100%	0.0	100.00%	Putative methyltransferase	99.55	1.1e-13	Enzimas diversas
M1132	91	hypothetical protein NDM1_gp0018 [Acanthamoeba polyphaga minnavirus]	100%	2.00E-62	100.00%	Rubredoxin-like fold containing protein	83.85	0.7	Outras proteínas
M1133	143	hypothetical protein [Acanthamoeba castellanii mannavirus]	100%	4.00E-101	100.00%	sem hit	-	-	Proteína hipotética
M1134	137	poly(A) polymerase small subunit [Megavirus chilensis]	99%	4.00E-80	90.58%	Poly A polymerase regulatory subunit	99.95	5.1e-28	Transcrição
M1135	163	Poly A polymerase regulatory subunit [Acanthamoeba castellanii mannavirus]	100%	3.00E-114	100.00%	Poly A polymerase regulatory subunit	100	1.1e-40	Transcrição
M1136	213	hypothetical protein [Acanthamoeba castellanii mannavirus]	81%	3.00E-122	100.00%	sem hit	-	-	Proteína hipotética
M1137	142	hypothetical protein NDM1_gp0015 [Acanthamoeba polyphaga minnavirus]	80%	7.00E-74	100.00%	sem hit	-	-	Proteína hipotética
M1138	425	hypothetical protein NDM1_gp0014 [Acanthamoeba polyphaga minnavirus]	100%	0.0	99.76%	sem hit	-	-	Proteína hipotética
M1139	91	putative replication origin-binding protein [Hirudovirus strain Sangsue]	100%	4.00E-57	98.89%	sem hit	-	-	Replicação de DNA
M1140	84	hypothetical protein NDM1_gp0013 [Acanthamoeba polyphaga minnavirus]	100%	1.00E-55	100.00%	sem hit	-	-	Proteína hipotética

Proteína	Tamanho (aa)	Best hit no BLASTp	Query cover	e-value	Id. %	Best hit no Hlpyred	Prob. %	e-value	Categoria
M1141	232	Uncharacterized protein R11 [Acanthamoeba polyphaga mimivirus]	100%	1,00E-166	100.00%	sem hit	-	-	Proteína hipotética
M1142	85	uncharacterized Bro-N domain-containing protein [Acanthamoeba polyphaga mimivirus]	100%	1,00E-53	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M1143	202	T-orf172 domain-containing protein [Acanthamoeba castellanii manavirius]	100%	8,00E-140	99.30%	sem hit	-	-	Repetições/dominios com funções diversas
M1144	376	hypothetical protein MIML_gp0009 [Acanthamoeba polyphaga mimivirus]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1145	1052	Helicase [Acanthamoeba castellanii manavirius]	100%	0.0	100.00%	Putative DEAD/DEAH box helicase	99.89	7.9e-21	Enzimas diversas
M1146	155	hypothetical protein MIML_gp0007 [Acanthamoeba polyphaga mimivirus]	100%	3,00E-108	100.00%	sem hit	-	-	Proteína hipotética
M1147	218	hypothetical protein MIML_gp0006 [Acanthamoeba polyphaga mimivirus]	100%	2,00E-157	100.00%	sem hit	-	-	Proteína hipotética
M1148	393	hypothetical protein [Povai lake negavirus]	96%	0.0	78.42%	sem hit	-	-	Proteína hipotética
M1149	404	KiIA N-terminal domain, NIR/P28 DNA binding protein [Acanthamoeba polyphaga lentillevirus]	92%	0.0	100.00%	KiIA-N domain-containing protein	97.73	0.00037	Repetições/dominios com funções diversas
M1150	178	KiIA N-family protein [Acanthamoeba polyphaga mimivirus]	100%	2,00E-119	96.63%	KiIA-N domain-containing protein	98.18	0.000037	Repetições/dominios com funções diversas
M1151	282	KiIA N-terminal domain, NIR/P28 DNA binding protein [Acanthamoeba polyphaga lentillevirus]	100%	0.0	100.00%	KiIA-N domain-containing protein	96.65	0.012	Repetições/dominios com funções diversas
M1152	614	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	0.0	100.00%	sem hit	-	-	Proteína hipotética
M1153	184	hypothetical protein MIML_gp0010 [Acanthamoeba polyphaga mimivirus]	99%	3,00E-126	100.00%	sem hit	-	-	Proteína hipotética
M1154	55	hypothetical protein vs_R309 [Acanthamoeba polyphaga lentillevirus]	100%	6,00E-27	100.00%	sem hit	-	-	Proteína hipotética
M1155	133	putative Bro-N domain-containing protein [Hirudovirus strain Sangsue]	93%	6,00E-84	99.20%	sem hit	-	-	Repetições/dominios com funções diversas
M1156	93	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	3,00E-61	100.00%	sem hit	-	-	Proteína hipotética
M1157	1052	Helicase [Acanthamoeba castellanii manavirius]	100%	0.0	100.00%	Putative DEAD/DEAH box helicase	99.86	2.6e-19	Enzimas diversas
M1158	155	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	1,00E-105	100.00%	sem hit	-	-	Proteína hipotética
M1159	125	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	9,00E-77	100.00%	sem hit	-	-	Proteína hipotética
M1160	311	BRO-N domain containing protein [Acanthamoeba castellanii manavirius]	100%	0.0	100.00%	sem hit	-	-	Repetições/dominios com funções diversas
M1161	312	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	0.0	99.04%	DNA primase small subunit	100	9.1e-39	Replacção de DNA
M1162	147	hypothetical protein [Acanthamoeba castellanii manavirius]	100%	5,00E-101	100.00%	sem hit	-	-	Proteína hipotética
M1163	406	hypothetical protein MIML_gp0009 [Acanthamoeba polyphaga mimivirus]	99%	1,00E-104	45.23%	sem hit	-	-	Proteína hipotética
M1164	187	putative KiIA-n domain-containing protein [Acanthamoeba polyphaga mimivirus]	13%	1,00E-04	100.00%	sem hit	-	-	Repetições/dominios com funções diversas

Anexo 3 – *Best hits* para proteínas dos transposirons já conhecidos anteriormente, obtidos no *HHpred*.

Transposirion do manuvirus			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
tv_R1	AEY99255.1	201	Proteína hipotética
tv_L2	AEY99254.1	140	Proteína hipotética
tv_L3	AEY99255.1	136	Proteína hipotética
tv_L4	AEY99256.1	112	Proteína hipotética
tv_L5	AEY99257.1	999	Proteína com domínio de helicase e integrase de ligação de zinco
tv_R6	AEY99258.1	89	Proteína com domínio de dedo de zinco tipo C2H2
Transposirion do lentilivírus			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
tv_R1	AEY99247.1	201	Proteína hipotética
tv_L2	AEY99248.1	140	Proteína hipotética
tv_L3	AEY99249.1	136	Proteína hipotética
tv_L4	AEY99250.1	112	Proteína hipotética
tv_L5	AEY99251.1	1027	Proteína com domínio de helicase e integrase de ligação de zinco
tv_R6	AEY99252.1	89	Proteína com domínio de dedo de zinco tipo C2H2
Transposirion do Mounovirus monte			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
tv_R1	AEY99259.1	144	Proteína hipotética
tv_L2	AEY99260.1	178	Proteína hipotética
tv_L3	AEY99261.1	103	Proteína hipotética
tv_L4	AEY99262.1	149	Proteína hipotética
tv_L5	AEY99263.1	113	Proteína com domínio de dedo de zinco tipo C2H2
tv_L6	AEY99264.1	1023	Proteína com domínio de helicase e integrase de ligação de zinco
tv_R7	AEY99265.1	73	Proteína com domínio de dedo de zinco tipo C2H2
tv_L8	AEY99266.1	87	Proteína hipotética
Transposirion do Mounovirus australiensis			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
matv_1	AVL93333.1	133	Regulador da transcrição
matv_2	AVL93334.1	176	Proteína hipotética
matv_3	AVL93335.1	132	Proteína hipotética
matv_4	AVL93336.1	104	Proteína hipotética
matv_5	AVL93337.1	136	Proteína com dedo de zinco C2H2
matv_6	AVL93338.1	1027	DNA helicase
matv_7	AVL93339.1	88	Proteína com dedo de zinco
matv_8	AVL93340.1	106	Proteína hipotética
Transposirion do negavirus courdo7*			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
tv_L1	AEY99242.1	148	Proteína hipotética
tv_L2	AEY99243.1	109	Proteína hipotética
tv_L3	AEY99244.1	79	Proteína com domínio de dedo de zinco tipo C2H2
tv_L4	AEY99245.1	1028	Proteína com domínio de helicase e integrase de ligação de zinco
tv_RS	AEY99246.1	89	Proteína com domínio de dedo de zinco tipo C2H2
Transposirion do megavirus vitis			
Proteína	Nº acesso (GenBank)	Tamanho (aa)	Anotação
mtv_v_1	AVL93326.1	105	Proteína hipotética
mtv_v_2	AVL93327.1	172	Proteína hipotética
mtv_v_3	AVL93328.1	116	Proteína hipotética
mtv_v_4	AVL93329.1	150	Proteína hipotética
mtv_v_5	AVL93330.1	167	Proteína com dedo de zinco C2H2
mtv_v_6	AVL93331.1	1028	DNA helicase
mtv_v_7	AVL93332.1	89	Proteína com dedo de zinco

Proteína	Hir HHpred	Probabilidade	e-value
tv_R1	Sem hit	-	-
tv_L2	Transmembrane protein: alpha-helical	91.44	1
tv_L3	Sem hit	-	-
tv_L4	Transposase (Uncharacterized protein Cg12762)	98.52	9.1e-7
tv_L5	EXOXYRIBONUCLEASE V ALPHA CHAIN; RECOMBINATION, HELICASE, NUCLEASE, Zinc finger homeobox protein	100	1.9e-31
tv_R6	Zinc finger homeobox protein	95.21	0.053

Proteína	Hir HHpred	Probabilidade	e-value
tv_R1	Sem hit	-	-
tv_L2	Transmembrane protein: alpha-helical	91.44	1
tv_L3	Sem hit	-	-
tv_L4	Transposase (Uncharacterized protein Cg12762)	98.52	9.1e-7
tv_L5	EXOXYRIBONUCLEASE V ALPHA CHAIN; RECOMBINATION, HELICASE, NUCLEASE, Zinc finger homeobox protein	100	1.1e-32
tv_R6	Zinc finger homeobox protein	95.21	0.053

Proteína	Hir HHpred	Probabilidade	e-value
tv_R1	Punative transcriptional regulator	98.05	0.000022
tv_L2	Sem hit	-	-
tv_L3	Sem hit	-	-
tv_L4	Sem hit	-	-
tv_L5	Zinc finger homeobox protein	90.24	0.47
tv_L6	EXOXYRIBONUCLEASE V ALPHA CHAIN; RECOMBINATION, HELICASE, NUCLEASE, zinc finger protein	100	2.00E-39
tv_R7	zinc finger protein	93.9	0.14
tv_L8	zinc hook motif and AAA domain	96.69	0.02

Proteína	Hir HHpred	Probabilidade	e-value
matv_1	TRANSCRIPTION REGULATION, DNA-BINDING, TETR FAMILY	99.02	3.00E-08
matv_2	Sem hit	-	-
matv_3	Sem hit	-	-
matv_4	Sem hit	-	-
matv_5	Zinc finger homeobox protein	91.99	0.2
matv_6	Exodeoxyribonuclease V, subunit RecD; alpha and beta protein, ATP-binding, Nucleotide-binding	100	1.00E-32
matv_7	Zinc finger homeobox protein	93.53	0.12
matv_8	Sem hit	-	-

Proteína	Hir HHpred	Probabilidade	e-value
tv_L1	Sem hit	-	-
tv_L2	Sem hit	-	-
tv_L3	Zinc finger homeobox protein	92.61	0.22
tv_L4	Exodeoxyribonuclease V, subunit RecD; alpha and beta protein, ATP-binding, Nucleotide-binding	99.98	6.8e-30
tv_RS	Zinc finger homeobox protein	95.64	0.027

Proteína	Hir HHpred	Probabilidade	e-value
mtv_v_1	Sem hit	-	-
mtv_v_2	Sem hit	-	-
mtv_v_3	Sem hit	-	-
mtv_v_4	Sem hit	-	-
mtv_v_5	Zinc finger homeobox protein	88.12	0.54
mtv_v_6	EXOXYRIBONUCLEASE V ALPHA CHAIN; RECOMBINATION, HELICASE, NUCLEASE, Zinc finger homeobox protein	100	5.8e-44
mtv_v_7	Zinc finger homeobox protein	95.21	0.059

Anexo 4 - Alinhamento (*BLASTp*) das proteínas do transpoviron do mimivírus argenteum contra o conjunto de proteínas dos outros transpovirons já descritos.

T1 - Proteína hipotética 1			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum	201	-	-	-	-
Transpoviron do lentilivirus - A	201	AEY99247.1	100%	4,00E-150	100.00%
Transpoviron do mamavirus - A	201	AEY99253.1	100%	4,00E-150	100.00%
Transpoviron do mousovirus monve - B	178	AEY99260.1	50%	6,00E-18	38.74%
Transpoviron do mousovirus australiensis - B	176	AVL93334.1	45%	2,00E-17	40.20%
Transpoviron do megavirus courdo7 - C	148	AEY99242.1	80%	3,00E-13	34.52%
Transpoviron do megavirus vitis - C	172	AVL93327.1	99%	4,00E-20	34.47%

T2- Proteína com domínio transmembrana (Tm helix)			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum	140	-	-	-	-
Transpoviron do lentilivirus - A	140	AEY99248.1	100%	4,00E-92	100.00%
Transpoviron do mamavirus - A	140	AEY99254.1	100%	5,00E-100	100.00%
Transpoviron do mousovirus monve - B	103	AEY99261.1	21%	0.036	30.00%
Transpoviron do mousovirus australiensis - B	-	-	-	-	-
Transpoviron do megavirus courdo7 - C	-	-	No significant similarity found		
Transpoviron do megavirus vitis - C	-	-	No significant similarity found		

T3- Proteína hipotética 2			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum - A	136	-	-	-	-
Transpoviron do lentilivirus - A	136	AEY99249.1	100%	1,00E-87	100.00%
Transpoviron do mamavirus - A	136	AEY99255.1	100%	2,00E-95	100.00%
Transpoviron do mousovirus monve - B	-	-	No significant similarity found		
Transpoviron do mousovirus australiensis - B	-	-	No significant similarity found		
Transpoviron do megavirus courdo7 - C	-	-	No significant similarity found		
Transpoviron do megavirus vitis - C	-	-	No significant similarity found		

T4- Transposase			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum -	112	-	-	-	-
Transpoviron do lentilivirus - A	112	AEY99250.1	100%	7,00E-73	100.00%
Transpoviron do mamavirus - A	112	AEY99256.1	100%	1,00E-80	100.00%
Transpoviron do mousovirus monve - B	144	AEY99259.1	52%	1,00E-09	37.29%
Transpoviron do mousovirus australiensis - B	133	AVL93333.1	45%	8,00E-09	45.10%
Transpoviron do megavirus courdo7 - C	-	-	No significant similarity found		
Transpoviron do megavirus vitis - C	-	-	No significant similarity found		

P5- Proteína com domínio de helicase			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum	999	-	-	-	-
Transpoviron do lentilivirus - A	1027	AEY99251.1	99%	0.0	100.00%
Transpoviron do mamavirus - A	999	AEY99257.1	100%	0.0	100.00%
Transpoviron do mousovirus monve - B	1023	AEY99264.1	99%	0.0	60.71%
Transpoviron do mousovirus australiensis - B	1027	AVL93338.1	99%	0.0	60.42%
Transpoviron do megavirus courdo7 - C	1028	AEY99245.1	99%	0.0	61.51%
Transpoviron do megavirus vitis - C	1028	AVL93331.1	99%	0.0	61.51%

P6- Proteína com dedo de zinco			Alinhamento <i>BLASTp</i>		
Transpoviron - clado	Tamanho (aa)	Nº de acesso (GenBank)	Query cover	e-value	Identidade
Transpoviron do mimivirus argenteum	89	-	-	-	-
Transpoviron do lentilivirus - A	89	AEY99252.1	100%	7,00E-57	100.00%
Transpoviron do mamavirus - A	89	AEY99258.1	100%	1,00E-64	100.00%
Transpoviron do mousovirus monve - B	73	AEY99265.1	65%	2,00E-10	36.21%
	113	AEY99263.1	69%	8,00E-07	33.33%
	87	AEY99266.1	47%	1,00E-05	33.33%
Transpoviron do mousovirus australiensis - B	88	AVL93339.1	35%	2,00E-09	50.00%
	136	AVL93337.1	50%	8,00E-07	41.30%
Transpoviron do megavirus courdo7 - C	89	AEY99246.1	87%	3,00E-17	44.30%
	79	AEY99244.1	32%	2,00E-06	48.28%
Transpoviron do megavirus vitis - C	89	AVL93332.1	87%	4,00E-18	44.30%
	167	AVL93330.1	32%	3,00E-06	48.28%