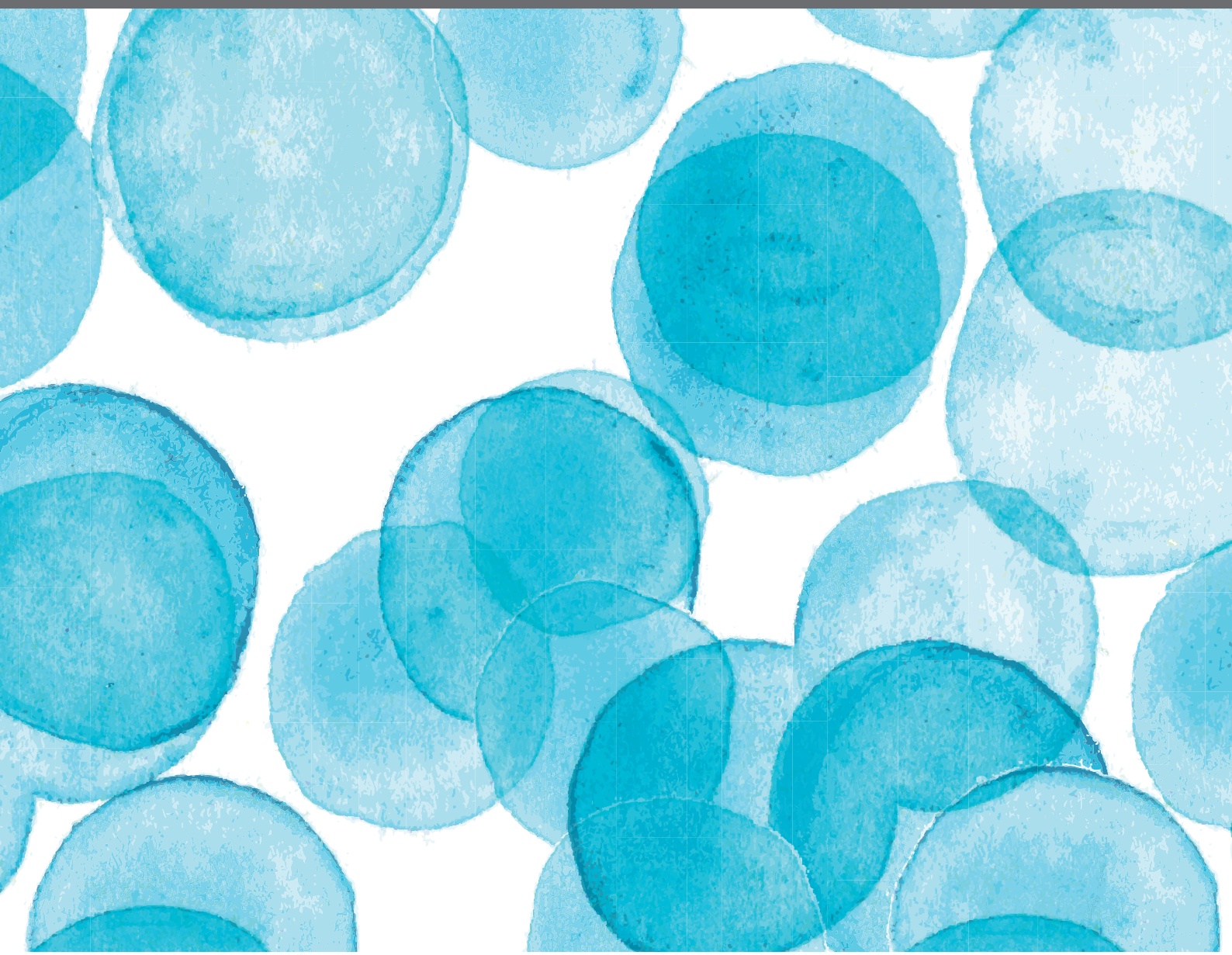
UPDATES ON LARGE AND GIANT DNA VIRUSES

EDITED BY: Jônatas Santos Abrahão and Bernard La Scola

PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88963-016-5
DOI 10.3389/978-2-88963-016-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

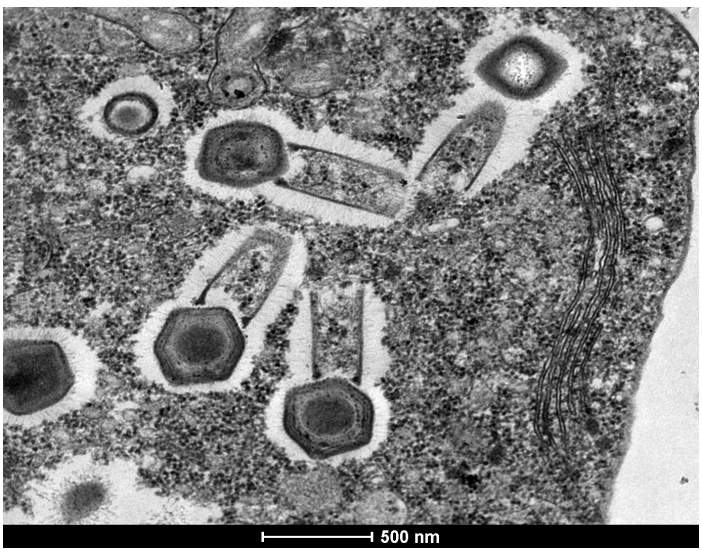
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

UPDATES ON LARGE AND GIANT DNA VIRUSES

Topic Editors:
Jônatas Santos Abrahão, Universidade Federal de Minas Gerais, Brazil
Bernard La Scola, Institut Hospitalo-Universitaire, France



Particles of Tupanvirus soda lake (TPVsl) in Vermoameba vermiformis under transmission electron microscopy (TEM).
Image: Microscopy Center, UFMG, Brazil/Jonatas Abrahao

Citation: Abrahão, J. S., Scola, B. L., eds. (2019). Updates on Large and Giant DNA Viruses. Lausanne: Frontiers Media. doi: 10.3389/978-2-88963-016-5

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 10 July 2019
doi: 10.3389/fmicb.2019.01608

Editorial: Large and Giant DNA Viruses

Jônatas Abrahão[1]* *and Bernard La Scola*[2,3]*

[1] *Laboratório de Vírus, Instituto de Ciências Biológicas, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil,* [2] *Microbes, Evolution, Phylogeny and Infection (MEΦI), Aix-Marseille Université UM63, Institut de Recherche pour le Développement IRD 198, Assistance Publique—Hôpitaux de Marseille (AP-HM), Marseille, France,* [3] *Institut Hospitalo-Universitaire (IHU)—Méditerranée Infection, Marseille, France*

Keywords: giant virus, large viruses, DNA virus, NCLDVs, evolution, pathogenesis

Editorial on the Research Topic

Large and Giant DNA Viruses

Since the seminal studies involving bacteriophages, the DNA viruses have fascinated the scientific community. DNA viruses were essential not only for the understanding of viral biological process, but also were a fundamental tool for the discovery and expanding knowledge related to cellular processes, such as transcription, translation, DNA repair, glycosylation and others. DNA viruses were also important characters during human history and evolution. The lethal and terrifying infection caused by a DNA virus, the smallpox disease caused by the variola virus, shaped and defined patterns of human migration, societies' interactions and raised innovative public health measures. In recent decades, some DNA viruses have been used as tools for heterologous protein expression and delivery, improving the field of vaccinology and diagnosis. In addition, some years ago, the discovery of the first mimiviruses shed new light on the study of DNA viruses field. Since then, many interdisciplinary studies, from distinct research groups, revealed breath-taking and controversial data regarding the origins, evolution and ecology of large and giant viruses. In this Research Topic, we received contributions from several colleagues on a broad range of topics related to large and giant DNA viruses.

Rodrigues et al. present a comprehensive meta-analysis of the currently known virosphere. In this study, it is crystal-clear that a substantial amount of knowledge on virology was obtained based on anthropocentric interests. The organisms with more viruses associated are human beings, plant crops, and domestic animals, revealing a huge gap on studies focused on viruses infecting species not related to humans. Contradicting this trend, we received many contributions on the discovery and biology of giant viruses that infect amoeba. A new and remarkable giant virus called Orpheovirus is described by Andreani et al. Orpheovirus is able to infect Vermoameba vermiformis and, with a genome exceeding 1.3 Mb and virions up to 1,300 nm in diameter, is one of the largest viruses described so far. Phylogenetic analysis provided evidence for a relationship between Orpheovirus and Pithovirus, however, some genetic characteristics revealed this new giant virus's divergent, independent evolution.

Silva et al. present an analysis of tupanvirus in *Vermoameba vermiformis*. Tupanvirus, a tailed giant virus, is the first to our knowledge that is able to infect more than one amoeba genus. In this paper, we learn that tupanvirus replication cycle in *V. vermiformis* is similar to tupanvirus cycle in *Acanthamoeba castellanii*. Outstanding scanning and electron microscopy images revealed fundamental steps of the cycle, including entry, factory formation, particle morphogenesis (including viral particle tail sprouting from factory), cell lysis and defective particles. The host-range of Marseillevirus (a virus discovered associated to *Acantamoeba*) was also explored by Aherfi et al. In this paper, the authors presented experimental inoculation of Marseillevirus in rats

OPEN ACCESS

Edited by:
Steven M. Short,
University of Toronto Mississauga, Canada

Reviewed by:
Jim L. Van Etten,
University of Nebraska-Lincoln, United States
Steven Wilhelm,
The University of Tennessee, Knoxville, United States

***Correspondence:**
Jônatas Abrahão
jonatas.abrahao@gmail.com
Bernard La Scola
bernard.la-scola@univ-amu.fr

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

Received: *24 April 2019*
Accepted: *26 June 2019*
Published: *10 July 2019*

Citation:
Abrahão J and La Scola B (2019) Editorial: Large and Giant DNA Viruses. Front. Microbiol. 10:1608. doi: 10.3389/fmicb.2019.01608

and mice models. Results revealed that, regardless the infection pathway utilized, Marseillevirus can be detected long-term in some organs, raising questions about the infective potential of this virus or a close relative in humans as suspected from cases of adenitis and lymphoma.

Evolutionary studies on giant virus were also explored in our Research Topic. Colson et al. performed a comprehensive study on the origins and ancestrality of giant viruses. By using phylogenetic and phenetic analyses, and the study of protein folding, to compare giant viruses and selected bacteria, archaea and eukaryota, the authors used their results to support the idea that giant viruses may cluster in a 4th branch of life, called 4th TRUC (for "Things Resisting Uncompleted Classifications"). This paper fuels the continuing (and perhaps controversial) debate on the origin of giant viruses. Chelkha et al. presented a phylogenomic study of the *Acanthameba polyphaga* draft genome, revealing more than 300 genes matching with viruses, including Pandoravirus, mimiviruses, Mollivirus sibericum, marseilleviruses, and Pithovirus sibericum. In a few case, genes seem to have been transferred from giant viruses to *A. polyphaga*, whereas in most of the cases the origins of those genes are equivocal. Assis et al. presented the genome characterization of the first two mimivirus of lineage C isolated in Brazil, called Mimivirus gilmour (MVGM) and Mimivirus golden (MVGD). In addition, the authors analyzed the pangenome of viruses belonging to *Mimivirus* genera, highlighting that discovery of new mimivirus isolates still contribute to the expansion of the pangenome and the consolidation of the core gene set. Aherfi et al. reported the isolation of three new Pandoravirus isolates, namely P. massiliensis, P. braziliensis, and P. pampulha. The authors presented an in-depth characterization of those isolates, including transcriptomics and genomics. In addition, the proteomics of P. massiliensis was described. The pangenome of the putative Pandoraviridae family was presented, revealing a large open pangenome and a small core genome. Louazani et al. analyzed the transcriptome of Faustovirus E12, presenting unexpected and complex splicing of the capsid gene. A total of 13 exons have been identified for the major capsid protein gene, including canonical and non-canonical splicing sites.

We also gathered new insights from papers focused on poxviruses. Cao et al. demonstrated the suppressive effect of resveratrol on vaccinia virus replication in various cell types. In this paper, authors suggest that resveratrol suppress the synthesis of viral DNA, affecting post-replicative gene expression. Szulc-Dabrowska et al. presented a comprehensive study on ectromelia virus, host immune response and viral evasion. Borges et al. presented serological evidence of silent (or possibly unreported) vaccinia virus exposure and disease in equids in southeast Brazil where the virus has been implicated in exanthematous outbreaks in cattle and humans.

The complex pathway of particle head assembly in the giant Salmonela phage SPN3US was explored by Ali et al. They presented data suggesting that a given prohead protease is able to cleave thousands of head proteins in just a few minutes to facilitate a major remodeling of the prohead prior to DNA packaging, impacting on viral assembly, final structure, composition and genome length. Huang et al. showed that the ubiquitin-proteasome system is important for replication of Singapore Grouper Iridovirus. Interestingly, several genes related to the ubiquitin-proteasome system were up/down-regulated during virus infection, and ubiquitin-proteasome system destruction impaired virus replication. Zhou et al. presented a new platform for genetic editing of Pseudoravies virus. The authors described the utilization of fosmid libraries for rapid generation of recombinant viruses. Xu et al. told us about the development of the recombinant immunotoxin called BoScFv-PE38, which has specific binding affinity for Bovine herpesvirus 1 glycoprotein D. They demonstrated that BoScFv-PE38 is internalized into MDBK cells compartments that inhibit BoHV-1 replication. Therefore, BoScFv-PE38 can potentially be employed as a therapeutic agent for the treatment of BoHV-1 infection. Finally, Zhang et al. presented data obtained *in vivo* suggesting that infection of chickens by infectious anemia virus can impair vaccinal immunity against Marek's disease.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank all the contributors of this Research Topic and we wish you all a good reading.

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

frontiers in Microbiology

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

An Anthropocentric View of the Virosphere-Host Relationship

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

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

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

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

OPEN ACCESS

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

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

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

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

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

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

INTRODUCTION

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

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

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

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

MATERIALS AND METHODS

Dataset Preparation and Selection Criteria

Virosphere and Hosts

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

Viral Diversity

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

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

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

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

Human Viruses and Viral Tropism

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

Construction of Networks

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

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

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

RESULTS AND DISCUSSION

The Known Viruses Have a Very Narrow Host Range

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

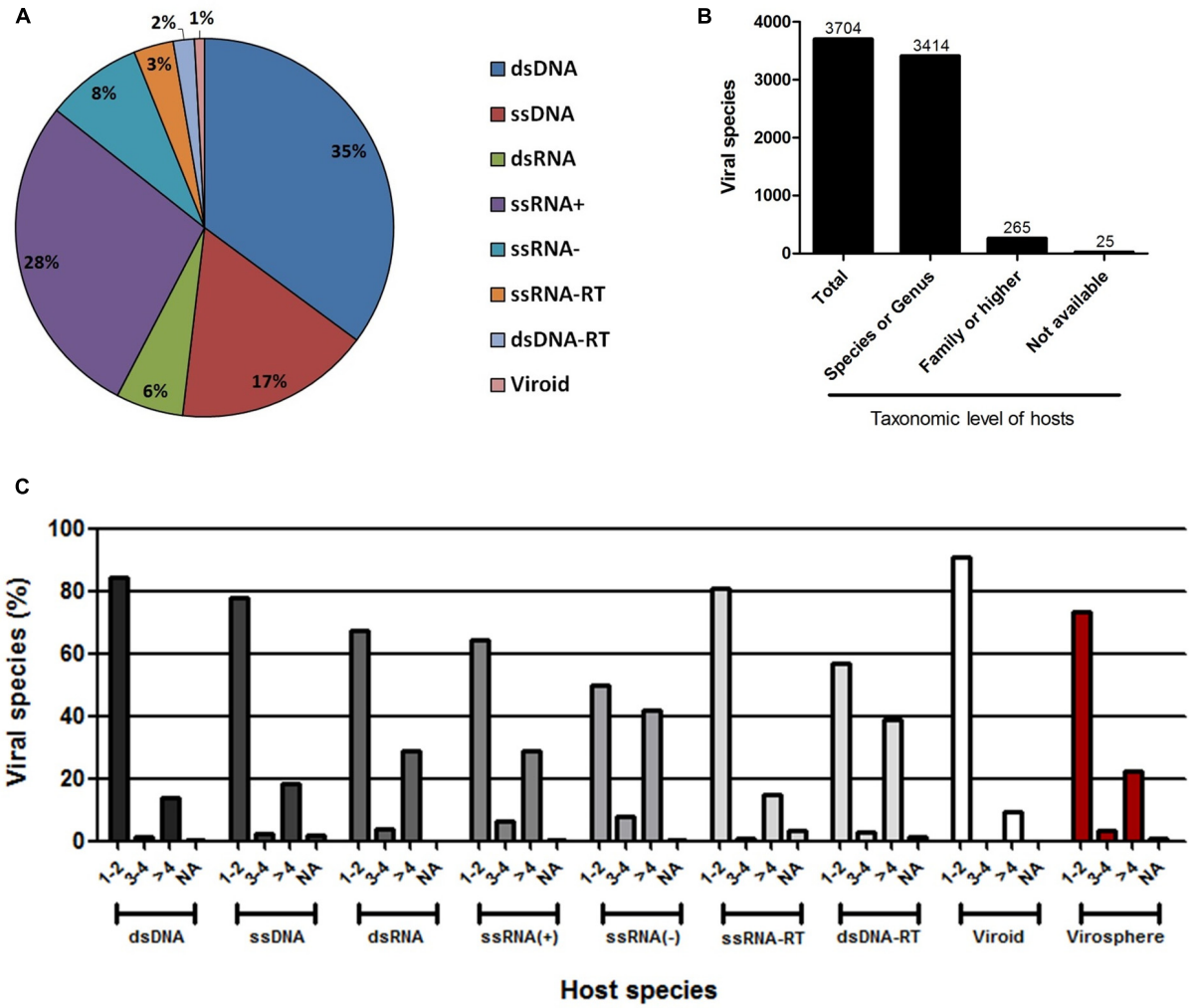


FIGURE 1 | Host range of the known virosphere. **(A)** Pie chart showing the distribution of the viruses recognized by the International Committee on Taxonomy of Viruses (ICTV) according to Baltimore's classification. **(B)** Taxonomic level of the hosts associated to the known viruses. More than 90%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

An Anthropocentric View of the Known Virosphere

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

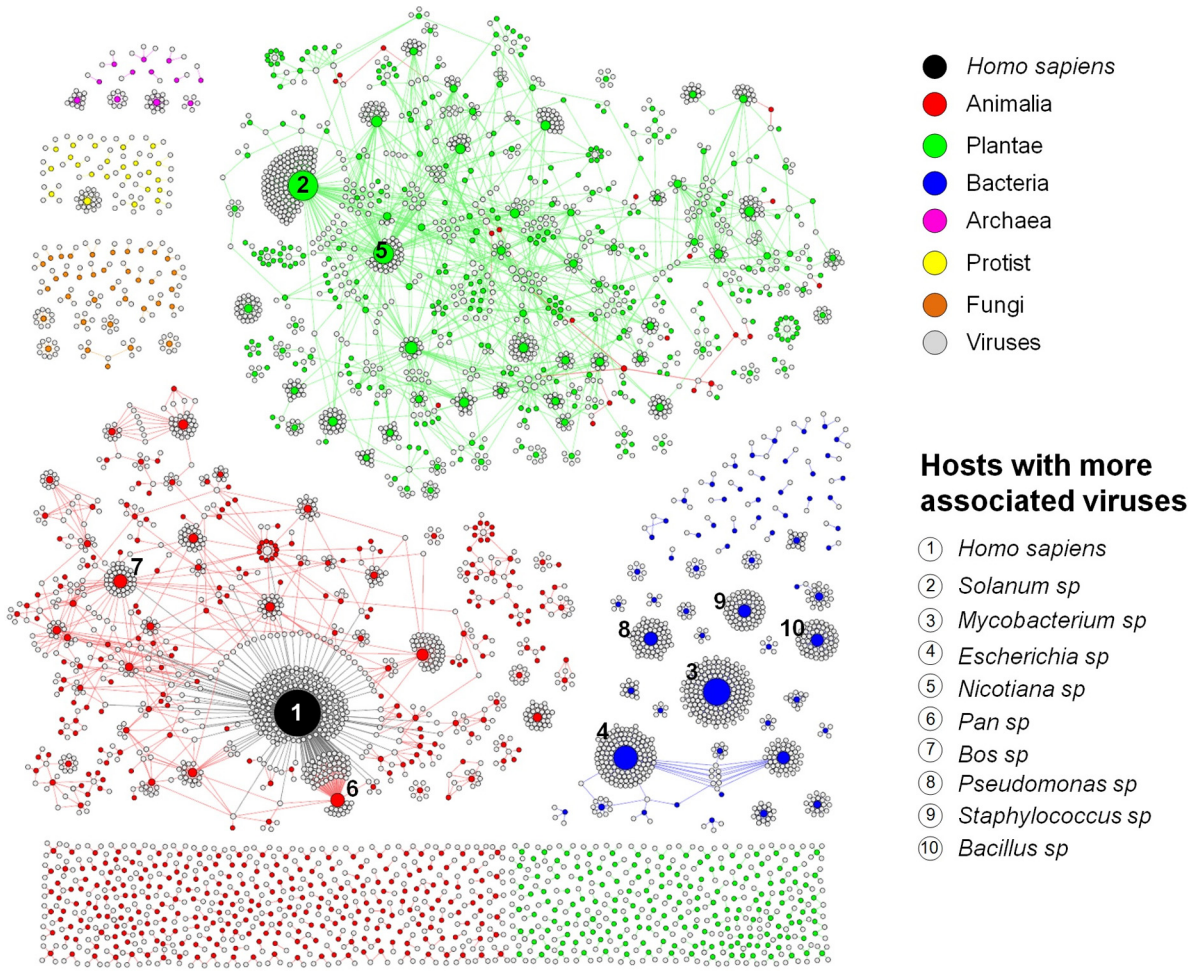


FIGURE 2 | Virus–host interaction network. Bipartite network graph showing a spatially connected network among viruses and hosts, a reflection of our limited knowledge about the viruses and their hosts. Each node represents a virus (gray), or a host genus, classified according to the taxonomic group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

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

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

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

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

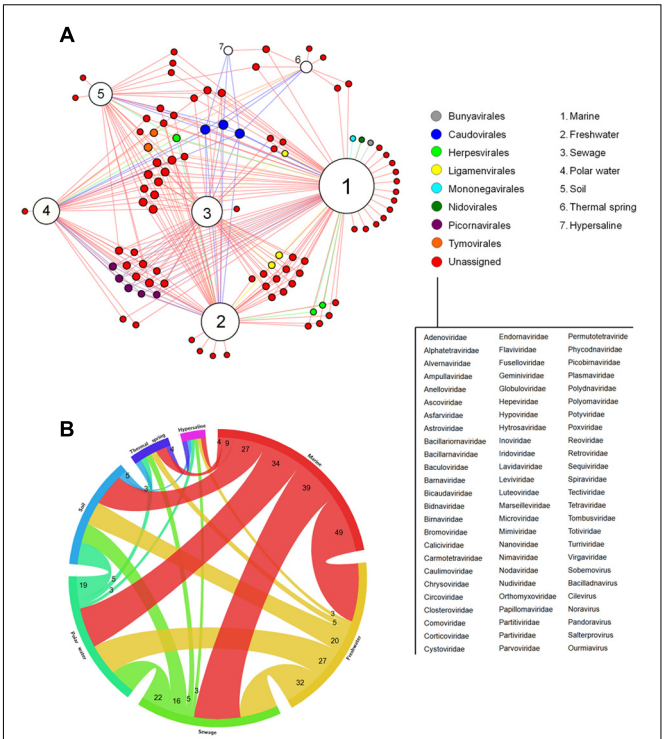


FIGURE 3 | Viral diversity on Earth. **(A)** Network graph showing the viral groups detected in different environments by metaviromic analysis. Each node represents an environment (white) or viral groups (families or known viral genus – colored nodes) classified according to the orders formally recognized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

Viral Diversity on Earth

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

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

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

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

Human-Associated Viruses and Viral Tropism

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

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

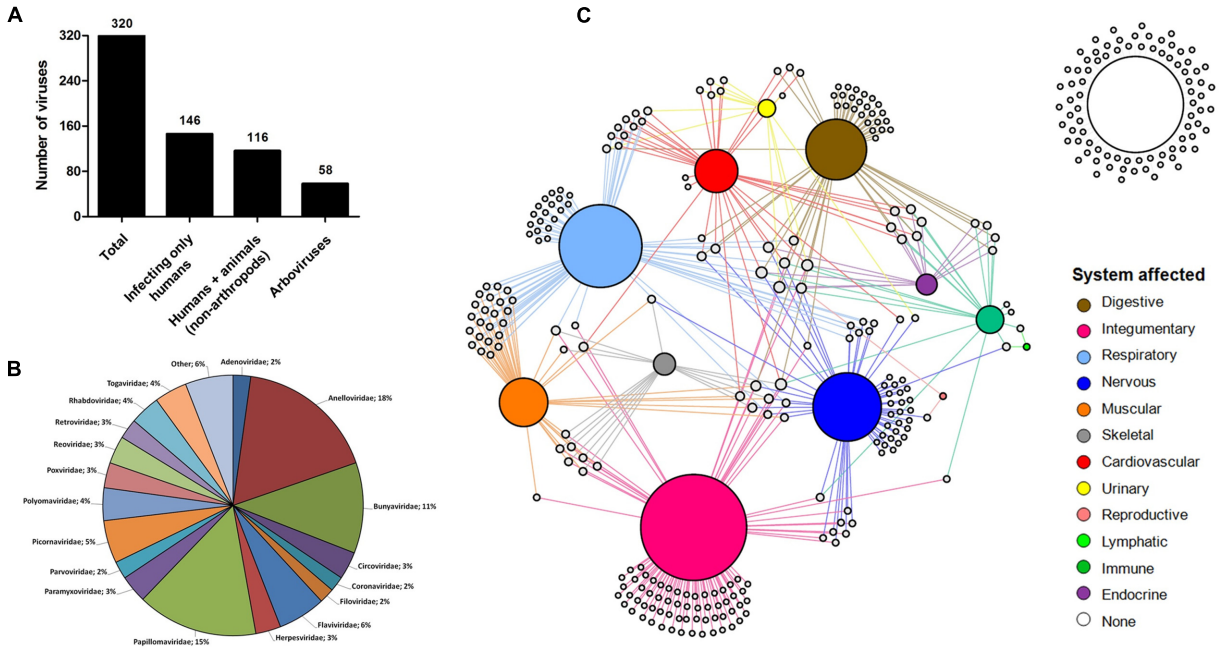


FIGURE 4 | Human viruses and affected systems. **(A)** Human-affecting viruses divided among infecting only humans, infecting humans and other mammals, and arboviruses. **(B)** Pie chart showing the classification of the viruses. A total of 27 groups are represented in the chart. Others: *Deltavirus,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

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

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

CONCLUSION

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

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

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

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPEMIG (Fundação de Amparo à Pesquisa do estado de Minas Gerais).

ACKNOWLEDGMENTS

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

SUPPLEMENTARY MATERIAL

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

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

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

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

TABLE S1 | Viruses and their hosts.

TABLE S2 | Technical information of metaviromic works.

TABLE S3 | Human-infecting viruses and other animals.

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

REFERENCES

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

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

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

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21848–21853. doi: 10.1073/pnas.0911354106

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Woolhouse, M., Scott, F., Hudson, Z., Howey, R., and Chase-Topping, M. (2012). Human viruses: discovery and emergence. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 2864–2871. doi: 10.1098/rstb.2011.0354

Zhang, C., Krause, D. J., and Whitaker, R. J. (2013). *Sulfolobus islandicus*: a model system for evolutionary genomics. *Biochem. Soc. Trans.* 41, 458–462. doi: 10.1042/BST20120338

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 22 January 2018
doi: 10.3389/fmicb.2017.02643

Orpheovirus IHUMI-LCC2: A New Virus among the Giant Viruses

*Julien Andreani[1], Jacques Y. B. Khalil[1,2], Emeline Baptiste[1], Issam Hasni[1], Caroline Michelle[1], Didier Raoult[1], Anthony Levasseur[1] and Bernard La Scola[1]**

[1] Aix Marseille Université, IRD, APHM, MEPHI, IHU-Méditerranée Infection, Marseille, France, [2] Centre National de la Recherche Scientifique, Marseille, France

Giant viruses continue to invade the world of virology, in gigantic genome sizes and various particles shapes. Strains discoveries and metagenomic studies make it possible to reveal the complexity of these microorganisms, their origins, ecosystems and putative roles. We isolated from a rat stool sample a new giant virus "Orpheovirus IHUMI-LCC2," using *Vermamoeba vermiformis* as host cell. In this paper, we describe the main genomic features and replicative cycle of Orpheovirus IHUMI-LCC2. It possesses a circular genome exceeding 1.4 Megabases with 25% G+C content and ovoidal-shaped particles ranging from 900 to 1300 nm. Particles are closed by at least one thick membrane in a single ostiole-like shape in their apex. Phylogenetic analysis and the reciprocal best hit for Orpheovirus show a connection to the proposed *Pithoviridae* family. However, some genomic characteristics bear witness to a completely divergent evolution for Orpheovirus IHUMI-LCC2 when compared to Cedratviruses or Pithoviruses.

Keywords: Orpheovirus, Cedratvirus, Pithovirus, *Vermamoeba vermiformis*, giant viruses, NCLDV, Orpheoviridae, *Pithoviridae*

OPEN ACCESS

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

Reviewed by:
Hiroyuki Ogata, Kyoto University, Japan
Gwenael Piganeau, Observatoire Océanologique de Banyuls sur Mer, France

****Correspondence:***
Bernard La Scola
bernard.la-scola@univ-amu.fr

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

Received: *02 October 2017*
Accepted: *19 December 2017*
Published: *22 January 2018*

Citation:
Andreani J, Khalil JYB, Baptiste E, Hasni I, Michelle C, Raoult D, Levasseur A and La Scola B (2018) Orpheovirus IHUMI-LCC2: A New Virus among the Giant Viruses. Front. Microbiol. 8:2643. doi: 10.3389/fmicb.2017.02643

INTRODUCTION

'Giant viruses' is a name commonly given to all viruses which are characterized by a capsid or ovoid shape, a size larger than 0.2 μm and a genome containing more than approximately 200,000 base pairs. This term encompasses a monophyletic group of large double stranded DNA viruses known as the nucleo-cytoplasmic large DNA viruses (NCLDV) (Iyer et al., 2001). The discovery of many Mimiviruses (La Scola et al., 2003; Arslan et al., 2011), Marseilleviruses (Boyer et al., 2009; Dornas et al., 2016) and Pandoraviruses (Philippe et al., 2013; Antwerpen et al., 2015), broke the paradigm of the previously held definition of the frontier between prokaryote and viruses. The "Megavirales" order proposed by Colson et al. (2013) continues to expand to host new arrivals with the potential of replacing the current NCLDV families (Colson et al., 2013; Aherfi et al., 2016). All these viruses share fundamental genes, for example, the conserved five ancestral genes and some others established into clusters of orthologous genes named NCVOGs (Yutin et al., 2009). Their replicative strategies appear to have adapted through their own evolution, as is the case for *Pandoraviruses* or *Mollivirus sibericum* (Abergel et al., 2015; Colson et al., 2017). Major improvements in taxonomy would be needed to definitely classify viruses in their families and in the putative "Megavirales" order. Further investigations should be focused on their genome content, hosts, ecosystems, tropisms and infectivity in order to determine whether their evolution is expansive or reductive or if it happens in a more dynamic accordion-like pattern (Moreira and Brochier-Armanet, 2008; Filée, 2014, 2015; Yutin et al., 2014; Moreira and López-García, 2015).

For now, co-culture on amoeba remains the major tool for isolating giant viruses (Pagnier et al., 2013; Khalil et al., 2016a). We recently combined co-culture with flow cytometry to come up with a faster and more sensitive way of detecting, presumably identifying and purifying the causative agent of lysis (Khalil et al., 2016b, 2017). In 2013, *Pithovirus sibericum* was isolated from a 30,000-year-old sample in the Siberian permafrost, and was described as being the most elongated-ovoid shape currently known for a virus with a maximum length of 1.5 μm. Surprisingly, the circular genome size is "only" of 610,033 base pairs, which appears to be astonishing given their viral particle size. The genome of *P. sibericum* is delivered via a single cork. Two years after the description of Pithovirus, a modern one that we named "*Pithovirus massiliensis LC8*" (Levasseur et al., 2016a) was also isolated and displayed amazing and extreme genomic conservation regarding its ancestor *P. sibericum*, which enabled us to estimate a molecular clock about the evolution of Pithoviruses. Moreover, we recently described a new virus Cedratvirus A11 (Andreani et al., 2016) a possible new genus in the putative *Pithoviridae* family. This virus presented two corks, one at each extremity and a circular genome estimated at 589,068 base pairs. In addition, a new strain, close to Cedratvirus A11, known as *Cedratvirus lausannensis,* was recently isolated (Bertelli et al., 2017) with a genome size estimated at 575,161 base pairs. This latter appears to represent a fourth member of this new emerging family. Our isolated Faustovirus (Bou Khalil et al., 2016) and Pithovirus (Levasseur et al., 2016a) indeed came from the same sampling area. For this reasons and after successfully isolating these viruses, we decided to investigate the same location once again, 4 months later in order to search for the same isolates that could be circulating and to explore the relation to ecosystemic or environmental changes. The result of this work was a new isolate from rat stool sample, which we named Orpheovirus, the genome and replicative cycle features of which we describe in this paper.

MATERIALS AND METHODS

Sample Collection

Twelve different rat stools and nine water samples contaminated by proximity sewage were collected. Rat stools were taken from a dry place one meter from the water sample area. Samples were harvested in November 2015 in La Ciotat, France, at the same GPS location where *P. massiliensis* LC8 (Levasseur et al., 2016a) and Faustovirus LC9 samples had been collected (Bou Khalil et al., 2016; Cherif Louazani et al., 2017) (N43.181834, E5.614423).

Virus Isolation

Vermamoeba vermiformis stain CDC19 was used as cell support. The amoebas were harvested after 48 h of culture in homemade peptone yeast extract glucose medium (PYG) when a concentration of 1.10^6 amoebas/mL was reached. Cells were then rinsed twice in homemade page's amoeba saline (PAS) and pelleted at 700 × g for 10 min. The amoebas were then re-suspended in the starvation medium (Bou Khalil et al., 2016) at a concentration of 1.10^6 amoebas/mL. An antibiotic and antifungal mixture with vancomycin (10 μg/mL), ciprofloxacin (20 μg/mL), imipenem (10 μg/mL), and voriconazole (20 μg/mL) was added to the suspension in order to decrease or eliminate bacterial or fungal contamination. A cell suspension of 250 μL per well was then distributed onto a 48-well plate. The samples were then vortexed and 50 μL were added to each well. The rest of the wells served as negative controls by adding 50 μL of PAS. The plate was incubated at 30°C for 4 days in order to monitor any potential cytopathic effect. This co-culture was repeated twice in the same order. When confronted with a high degree of contamination detected in some wells, filtration using 1.2 μm syringe filter (Merck Millipore) was carried out and gentamycin (20 μg/mL) was added 24 h before the second plate of co-culture (sub-culture 1).

Viral Production and Purity Control

End-point dilution was performed in order to clone the virus before its production. To do so, we successively inoculated diluted viral supernatant on *V. vermiformis* at a dilution factor of 10. End point dilution was assessed for 5 days and the lysis was controlled by inverted microscopy.

For the production and purification processes, 14 infected flasks of 150 cm^2 (Corning®, Corning, NY, United States) were pelleted using the Beckman coulter® centrifuge Avanti® J-26 XP (Beckman, France) at 14,000 × *g* for 30 min (Andreani et al., 2016; Levasseur et al., 2016a). A 25% sucrose gradient was used for the final purification step. After finalizing production, we proceeded with genome sequencing.

Genome Sequencing

Genomic DNA was sequenced on the MiSeq Technology (Illumina Inc., San Diego, CA, United States) using the paired end and mate pair applications. The DNA was barcoded in order to be mixed with 11 other projects for the Nextera Mate Pair sample prep kit (Illumina) and with 16 other projects for the Nextera XT DNA sample prep kit (Illumina).

gDNA was quantified using a Qubit assay with the high sensitivity kit (Life Technologies, Carlsbad, CA, United States) to 131.3 ng/μl.

For the paired end library, dilution was performed requiring 1ng of each genome as input. The "tagmentation" step fragmented and tagged the DNA. Limited cycle PCR amplification (12 cycles) then completed the tag adapters and introduced dual-index barcodes. The library profile was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc., Santa Clara, CA, United States) with a DNA High sensitivity labchip and the fragment size was estimated to 1.5 kb. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, United States), the libraries were then normalized on specific beads according to the Nextera XT protocol (Illumina). Normalized libraries were pooled for MiSeq sequencing. Automated cluster generation and paired end sequencing with dual index reads were performed in a single 39-h run in 2 × 250-bp.

A total of 6.6 Gb of information was obtained from a 697,000 per mm^2 for the density cluster with a cluster passing quality control filters of 94.6% (12,733,000 passed filtered clusters). Within this run, the index representation for Orpheovirus IHUMI-LCC2 was determined to 12.93%. The 1,942,146 paired end reads were trimmed and filtered according to the read qualities.

The mate pair library was prepared with 1.5 μg of genomic DNA using the Nextera mate pair Illumina guide and two libraries were constructed. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc., Santa Clara, CA, United States) with a DNA 7500 labchip. The DNA fragments ranged from 1.5 kb to 11 kb with an optimal size at 6.57 and 2.89 kb, respectively. No size selection was performed and 600 and 117 ng, respectively, of tagged fragments were circularized.

The circularized DNA was mechanically sheared to small fragments with an optimal size of 1029 and 1253 bp, respectively, on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, United States).

The library profile was visualized using a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc., Santa Clara, CA, United States) and the final concentration libraries were measured at 5.13 and 5.4 nmol/l, respectively.

In each construction, the libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-h run in a 2 × 151-bp.

Total information of the two flowcells at 6.2 and 7.9 Gb was obtained from a 648,000 and 863,000 cluster density per mm^2 with a cluster passing quality control filters of 96.1 and 94% (12,144,000 and 15,627,000 passing filter paired reads). Within these runs, the index representation for Orpheovirus IHUMI-LCC2 was determined at 3.16 and 12.43%. The 725,401 and 1,942,196 paired reads were trimmed and assembled with the paired end reads.

Genome Assembly

Mate pair and paired-end reads were trimmed using CLC Genomics Workbench v7.5[1]. *De novo* assembly of all reads was conducted using 64-word size and 50 bubble size parameters. We obtained 20 scaffolds representing a total size 1,461,620 bp with an average coverage reads ranged from 423 to 551. In parallel, we used an A5 pipeline assembler (Tritt et al., 2012) with standard parameters on 3,884,384 raw reads (paired end reads) representing 621,103,741 nucleotides. We obtained one scaffold of 1,473,699 with a median coverage reads of 295 with a 10th percentile at a coverage of 226. However, two regions of repeats were not completely resolved. Blast alignments of the two different assembling strategies confirmed these two regions and also underlined a high degree of identity between the two methods of assembly (>99%). For these two regions on the A5 assembly, we used GapCloser (Luo et al., 2012) and GapFiller (Nadalin et al., 2012) to fill two gaps and obtained a final single scaffold of 1,473,573 base pairs.

Genome Alignments and Genome Organization

The MAUVE program (Darling et al., 2004) was used to align and determine nucleotide divergence between genomes. BLAST nucleotide online was used to generate dot plots to explore large repeats in the whole genome and in all specific coding sequences. Emboss Explorer was used online using the following different software programs: palindrome of a 200 maximum length[2], an e-inverted program, an equicktandem for a fast detection.

Genome Analysis

Gene prediction was computed using Genemarks software (Besemer et al., 2001). We deleted predicted proteins having a size less than 50 amino acids, and 85 predicted protein from 50 to 99 amino acids were detected by Phyre2 (Kelley et al., 2015) as having abnormal tri-dimensional folding and finally were discarded from our dataset. A Blast protein was performed against the non-redundant (nr) protein database (June 19, 2017). Annotation was performed using a combination of Interpro[3] version 63.0, a CD-search tool online (Marchler-Bauer and Bryant, 2004) and delta-blastp (Boratyn et al., 2012). Interpro detected 100 transmembrane domain-containing proteins, and with CD-search and delta-blastp they congruently identified domains in 443 proteins.

tRNA prediction was computed online[4] (Lowe and Chan, 2016) following different standard parameters successively with eukaryotes, archaea and bacteria. We identified orthologous and paralogous genes by using Proteinortho v5 (Lechner et al., 2011) with 60% coverage and 20% amino acid identity and an *e*-value of 10^{-2} as significance thresholds. Moreover, we generated pan-genomic tree on GET_HOMOLOGUES package (Contreras-Moreira and Vinuesa, 2013) using OrthoMCL algorithm with the standard parameters expected for the coverage and *e*-value. We choose 60% as minimum coverage in Blastp pairwise alignments and 1×10^{-2} as maximum *e*-value.

Genome Submission

Orpheovirus IHUMI-LCC2 is available in the EMBL-EBI database under accession number LT906555.

Phylogenetic Analysis

All phylogenetic analyses were conducted using the following procedures. Blastp was used to find close homologous proteins. Then, the MUSCLE program (Edgar, 2004) was used to align amino acid sequences. The FastTree program (Price et al., 2009) was computed with standard parameters using the maximum likelihood method with 1,000 bootstrap replicates and the Jones–Taylor–Thornton (JTT) model for amino acid substitution.

[1]http://www.clcbio.com/blog/clc-genomics-workbench-7-5/

[2]http://emboss.bioinformatics.nl/cgi-bin/emboss/palindrome

[3]https://www.ebi.ac.uk/interpro/

[4]http://trna.ucsc.edu/tRNAscan-SE/

Phylogenetic trees were then visualized using iTOL v3 online (Letunic and Bork, 2016).

RESULTS

Virus Isolation

Bacterial contamination is common in viral co-cultures when using stool and sewage samples, with the frequent presence of resistance to the antibiotics and anti-fungal mixtures used. For this, we used a classic mix of antibiotics notably, vancomycin, ciprofloxacin, and imipenem, as previously reported. However, we added gentamicin to our first sub-culture plate, 24 h before inoculating the new plate in order to eliminate resistant bacteria from the stool samples. After three passages on *V. vermiformis*, lysis referring to the cytopathic effect was detected in some wells. We performed negative staining on the supernatant of a rat stool in well LCC2 and observed particles with an elongated aspect (**Figure 1**), some of them appear to be irregular, with a concave shape compared to *Pandoravirus* and *Pithoviruses*. In contrast, the apex appears to be more similar to Pandoravirus. We named it Orpheovirus.

Replicative Cycle

The length of the Orpheovirus virions range from 900 to 1,100 nm ($N = 10$) with a maximum diameter of about 500 nm ($N = 13$). Some virions could reach 1,300 nm in length; this process was sometimes observed in the host cytoplasm. The cork does not seem to seal by a grid as opposed to Pithoviruses and Cedratviruses. We noticed shapes which were similar to the ostiole-like apex observed in Pandoraviruses (Philippe et al., 2013) with a diameter ranging from 70 to 80 nm ($N = 8$) obstructed by a thick membrane (**Figures 2F,I**). Nevertheless, in Pandoraviruses the tegument is composed of three layers, each measuring about 20 ∼ 25 nm. The Orpheovirus' particles presented a dark dense outside layer coated with short, sparse fibrils on their external surface (black arrow). This dark layer is followed by a medium dense space (white arrow) which is in direct contact with the thin inner hyperdense membrane surrounding the viral core cavity containing the nucleic acid (**Figure 2I**). Altogether, these layers measure ≈40 nm. The replicative cycle of Orpheovirus showed classical stages of infection and replication in amoeba. Briefly, the virus entry by phagocytosis is the start of the cycle, where particles escape the phagosomal process. DNA delivery occurs in the amoeba cytoplasm via the ostiole-like apex (**Figures 2A,B**). An eclipse phase takes place at 4 h post entry. Functional viral factories (**Figure 2C**) are well installed and detected around 14–16 h post-infection. Similar forms corresponding to early virion synthesis are also observed, as it is the case for Pithoviruses and Cedratviruses (**Figure 2D**). At 20 h post-infection, the host cells' cytoplasm is fully occupied by newly synthesized virions (**Figures 2E,G,H**). We were also able to detect viruses outside the amoeba due to cell burst or viruses exiting by exocytosis. Complete cell burst occurred between 24 and 38 h post-infection. This slow viral cycle is often observed in the case of *V. vermiformis* used as cell support, which is not the case when using *Acanthamoeba* spp. (Reteno et al., 2015; Andreani et al., 2017).

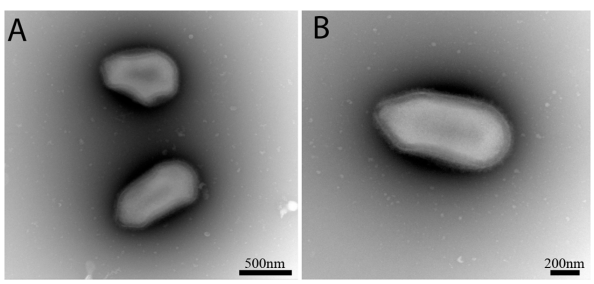


FIGURE 1 | Negative staining obtained from the co-culture supernatant of well LCC2. Scale bars are indicated on each panel. **(A)** Two specimens of Orpheovirus, the top one has a curious irregular form with highly concave forms. **(B)** Classic ovoid shape of Orpheovirus IHUMI-LCC2.

Orpheovirus: Main Genomic Characteristics

Orpheovirus has a circular genome estimated at 1,473,573 base pair (including 100 N due to an incomplete elucidate region in its genome) with a GC%-content established at around 25% (**Table 1**). A megablast or a simple blastn against the nr/nt nucleotide collection database revealed no match with other known giant viruses. Dot plots show various areas of repeats (Supplementary Figures S1–S3). We found 57 palindromic sequences, 1,527 tandem repeats and 832 inverted sequence candidates. The number of repeats explains the complexity observed during the genome assembly steps. A comparison with other giant viruses (Supplementary Table S1) showed an extremely high number of tandem repeats and inverted repeats for Orpheovirus.

1,512 genes were predicted but, following our method, 313 genes with an abnormal conformation already cited in the material and methods section were discarded. We only retained 1,199 genes, resulting in a coding density of around 66.4% (979,005 base pairs). This value is close to that of Pithoviruses but lower than that of Cedratvirus A11. A Blast against the nr database retrieved 509 matched proteins with at least one known protein (≈42.5% of all predicted proteins), and 690 unmatched, which are classified like ORFans genes (≈57.5% of all predicted genes). Of the 509 proteins, two had a hit with unclassified sequences, 148 had a best hit with a virus (≈12.3% of all conserved proteins), 176 with eukaryotes (≈14.7%), and 183 with prokaryotes (≈15.3%) (**Figure 3**). Regarding the 148 best hits with viruses, we observed 27 best hits with *P. sibericum*, 11 with *P. massiliensis*, 15 with Cedratvirus A11, 24 with Mimivirus A, B, and C lineages, and 18 with *Klosneuvirinae*. Hence, the highest best hit viral was obtained with the putative family *Pithoviridae* with 53 best hits, although the value was also important with *Mimiviridae* and associated extend family.

Despite this, the 57.5% of Orpheovirus' genes are ORFans, with an *e*-value cut-off of 10^{-2}. This could increase at ≈66% when we chose a more stringent cut-off value for the blastp at 10^{-5}. We found 343 genes that formed 167 clusters of

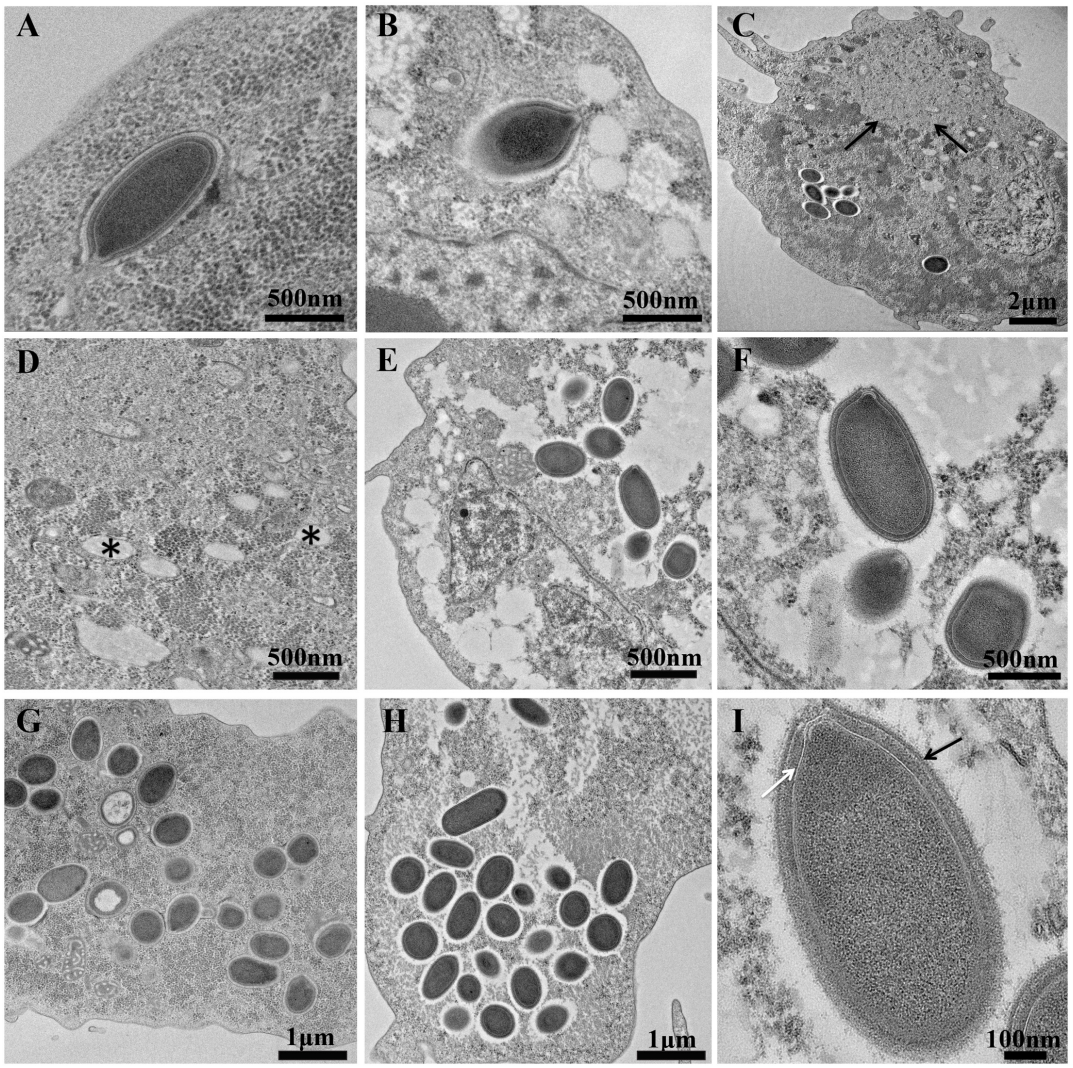


FIGURE 2 | Ultrathin sections of Orpheovirus's replicative cycle. Scale bars are indicated on each panel. **(A,B)** Represent viral entry at 2 h and 4 h post-infection. **(C)** Represents a section of *Vermamoeba vermiformis* 16 h post-infection, Black arrows delimitate the viral factory. **(D)** High magnification of **(C)** picture, (*) represents some curious vacuoles in contact with the viral factory in the cytoplasm. **(E–G)** Show some cytoplasms and new virus synthetized 20 h post-infection. **(H)** Accumulation of assembled virions at 20 h post-infection. **(I)** Single virion into the cytoplasm of *V. vermiformis* at 24 post-infection, Black arrow points to the external membrane and white arrow indicates the medium dense space.

TABLE 1 | Main genomic characteristics of Orpheovirus and other closely related viruses.

Virus	Orpheovirus IHUMI-LCC2	*Cedratvirus lausannensis* CRIB-75	Cedratvirus A11	*Pithovirus massiliensis* LC8	*Pithovirus sibericum*
Morphological features	Ovoid, single ostiole-like	Ovoid double corks	Ovoid double corks	Ovoid single cork	Ovoid, single cork
Genome size (bp)	1,473,573	575,161	589,068	686,015	610,033
GC content (%)	24.98	42.8	42.6	35.4	35.8
tRNA	0	0	0	0	0
Predicted proteins	1199	643	574	476	467
ORFans (%)[1]	≈66	≈45	≈35	N/A[2]	67.5
Coding density	66.4%	83%	78.5%	64%	69%

[1]ORFans are given at the moment of viral description with an e-value at 10^{-5}. [2]N/A, not applicable. All ORFans of P. sibericum are found in P. massiliensis (Levasseur et al., 2016a).

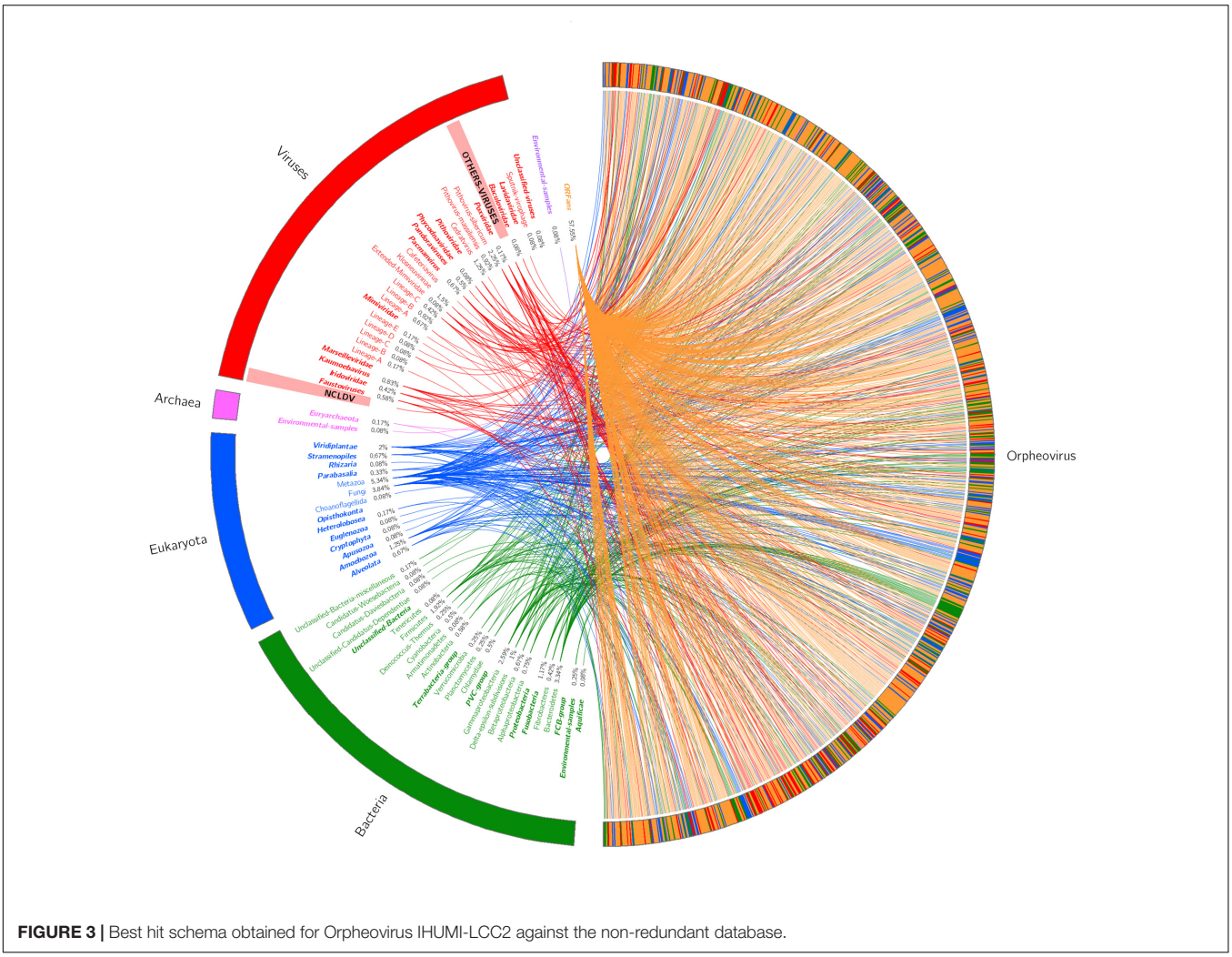


FIGURE 3 | Best hit schema obtained for Orpheovirus IHUMI-LCC2 against the non-redundant database.

paralogous genes in Proteinortho, regrouping few numbers of genes ranging from two to a maximum of four by cluster. Annotation of these clusters revealed predominant predicted proteins mainly as MORN-repeat (55 sequences), Ankyrin-repeat (37 sequences), and F-box domain-containing (81 sequences).

The Orpheovirus annotation presented translation system components as follows: eight aminoacyl tRNA synthetases (aaRS), four translation factors: three initiation factors and one release factor (Supplementary Table S2). Surprisingly, Orpheovirus didn't present any tRNA. We used the aminoacyl tRNA synthetase, which appeared to be a good way to distinguish and classify some lineages and to describe hypothetical common ancestor (Abrahão et al., 2017; Schulz et al., 2017). The Glycyl-tRNA synthetase of Orpheovirus was found to branch with the Asgard Glycyl-tRNA synthetase, and not with Catovirus CTV1 nor Klosneuvirus KNV1 homologs. This Asgard superphylum described by metagenomic studies seems to be a controversial bridge between prokaryotes and eukaryotes (Spang et al., 2015; Da Cunha et al., 2017; Zaremba-Niedzwiedzka et al., 2017). Regarding the phylogenetic analysis of each tRNA synthetase (Supplementary Figures S4–S11), we observed different patterns for amino acyl tRNA synthetase. While some are monophyletic with other described giant viruses, others appear to be polyphyletic resulting from potential lateral gene transfer.

Orpheovirus and Its Divergent Viral Neighborhood

First of all, we searched for five ancestral genes of NCLDV (Colson et al., 2013) encoding the major capsid protein, the helicase-primase (D5), the DNA polymerase elongation subunit family B, the DNA-packaging ATPase (A32), and the viral late transcription factor 3 VLTF3. Four of these genes were found with the exception of the A32-like packaging ATPase, which was absent in all four viruses (Legendre et al., 2014; Andreani et al., 2016). As reported for Cedratvirus A11, *P. massiliensis,* and *P. sibericum,* Orpheovirus presented two distinct RNA polymerase II subunit 1. Multiple ribonucleases such as Ribonuclease R, two Ribonuclease III and one ribonuclease HI were detected in Orpheovirus. Orpheovirus presented glycosyltransferase and numerous proteins involved in lipid pathways. We also identified two proteins presenting multiple fusion bacteria domains involved

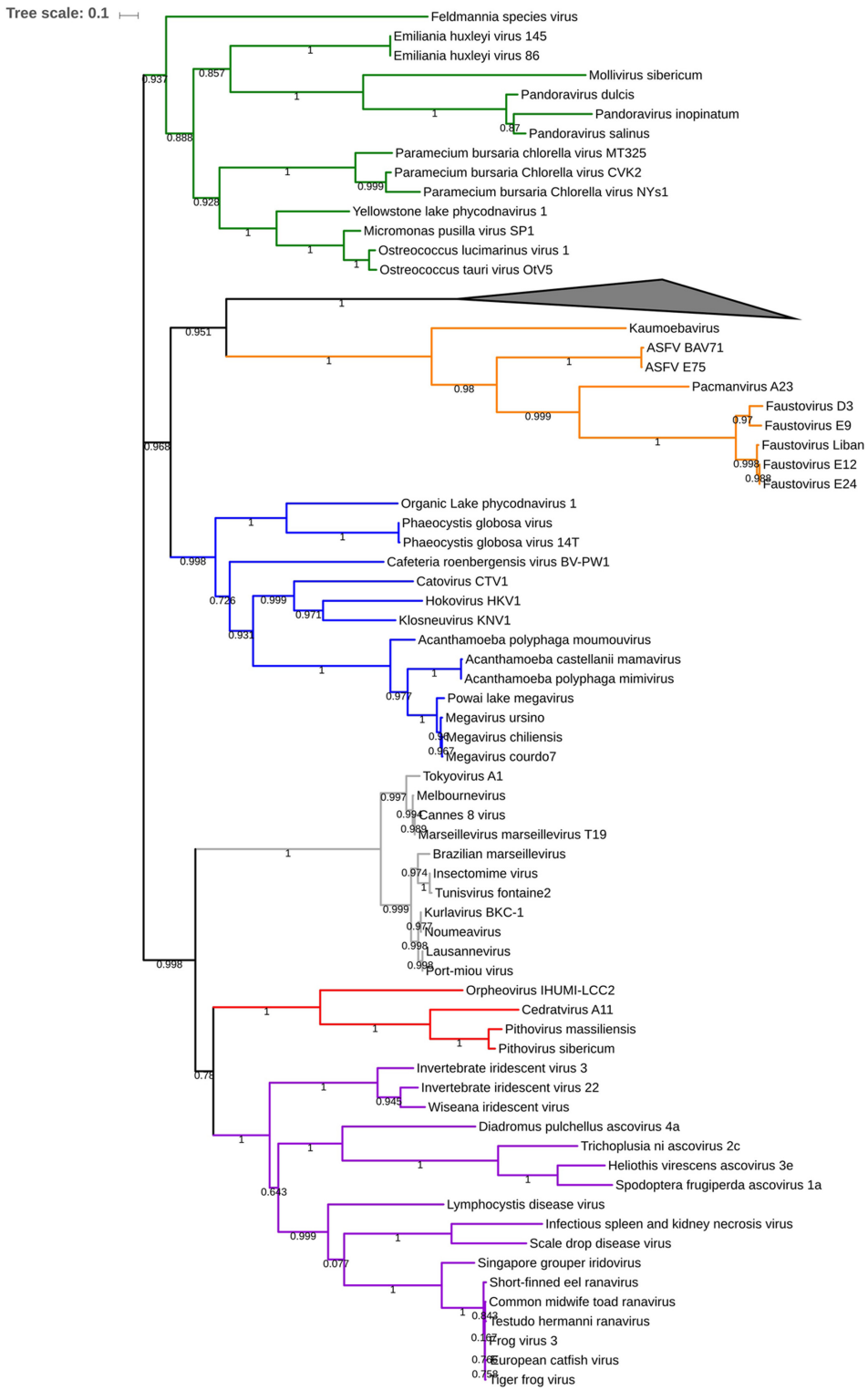


FIGURE 4 | Phylogenetic tree based on 84 DNA polymerase b protein of nucleo-cytoplasmic large DNA viruses (NCLDV). Branch values lower than a bootstrap value of 0.5 were deleted. Colors were assigned for different group of viruses: blue for Mimivirus and extended Mimiviridae; green for Pandoraviruses, *Mollivirus sibericum* and Phycodnaviridae; orange for groups of Asfarviridae, Faustoviruses, Pacmanvirus and Kaumoabevirus; gray for *Marseilleviridae*; red for Orpheovirus, Cedratvirus, and Pithoviruses and purple for *Asco-Iridoviridae*. The collapsed branch represented by a black triangle was used for 15 *Poxviridae* members. The corresponding alignment is available on Supplementary Data Sheet 2 visualized by automatic MView software (https://www.ebi.ac.uk/Tools/msa/mview/). 3,450 positions were used to build the tree.

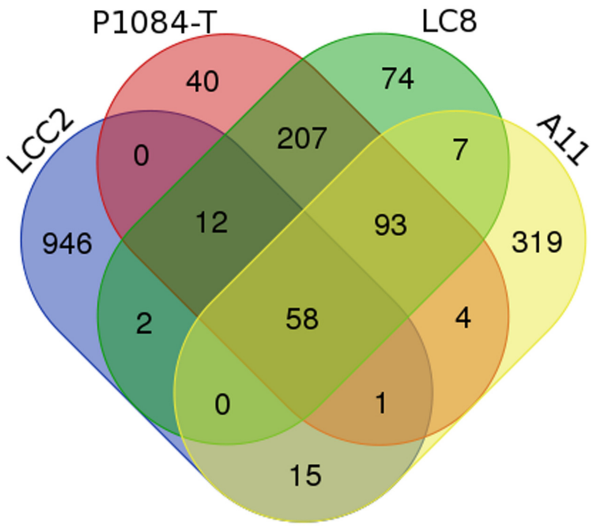


FIGURE 5 | Venn diagram comparison of best reciprocal hit obtained belonging to the predicted proteins of Orpheovirus, Cedratvirus, Pithoviruses. LCC2 represents Orpheovirus IHUMI-LCC2, P1084-T represents *Pithovirus sibericum*, LC8 represents *Pithovirus massiliensis* and A11 represents Cedratvirus A11. The diagram was created online http://bioinformatics.psb.ugent.be/webtools/Venn/.

in Riboflavin (Vitamin B2) biosynthesis. Indeed, we observed in ORPV_596 Tri-functional domains of "Di-Hydro-Folate-Reductase/deoxycytidylate deaminase/Riboflavin biosynthesis protein RibD" presenting a homology with Indivirus ILV1. And the second protein is ORPV_666, annotated like Tri-functional domains "3,4 dihydroxy-2-butanone 4-Phosphate synthase/GTP cyclohydrolase II/Lumazine synthase (RibA+RibB+RibH)" presenting homologies with Indivirus ILV1, *Bacillus subtilis,* and *Acanthamoeba castellanii strain Neff*. Vitreschak (2002) demonstrated that Riboflavin operon gene fusion is frequently found in bacteria (Vitreschak, 2002).

After that, phylogenetic analysis based on the DNA polymerase B protein, VLTF3 and RNA polymerase II subunit 1 showed deep branching with Cedratvirus and Pithoviruses (**Figure 4** and Supplementary Figures S12, S13). Moreover, 58 reciprocal best hit proteins were only shared between Orpheovirus IHUMI-LCC2, Cedratvirus A11, *P. sibericum P1084-T* and *P. massiliensis LC8*. In addition, 14 reciprocal best hit proteins were found to be shared between Orpheovirus and Pithoviruses (14+58) and 15 between Cedratvirus and Orpheovirus (15+58) (**Figure 5**), while Cedratvirus shared 151 proteins (58+93) with Pithoviruses. Meanwhile, 946 of 1,034 protein clusters (≈91.4%) are unique to Orpheovirus, 319 of 497 clusters (≈64.2%) to Cedratvirus A11, and 114 clusters of 543 (≈21%) to Pithoviruses. There were only two colinearity blocks and nine lines connecting Orpheovirus to other viruses (Supplementary Figure S14).

Following the discovery of this divergence between the four viruses, we decided to investigate Orpheovirus position in the "Megavirales" order further with the help of a parsimonious pan-genomic tree (Supplementary Figure S15). The long branch length observed for Klosneuvirus, *Pandoravirus inopinatum* and Orpheovirus is explained by the large genome size and by the number of predicted proteins compared to the close relative strains in the tree. These long branches could be a new common marker to explain the emergence of new viral family or lineages in the proposed "Megavirales" order. In the case of Orpheovirus, the pan-genomic analysis confirms this distant relation with the proposed *Pithoviridae* family.

Orpheovirus and Virophage: A Curious Homologous Sequence

Orpheovirus has a predicted gene a 434 amino acids protein that we called V21-like protein. This protein had homologs in Blastp, respectively, at 86% coverage, 21% identity with Sputnik virophage V21 protein (La Scola et al., 2008), and 85% coverage, 27% identity with Zamilon (Gaia et al., 2014). These two homologous proteins are annotated as hypothetical proteins, and showed no other homology using the blast strategy. However, HHpred online (Supplementary Data Sheet 3) and Phyre2 (Supplementary Figure S16) detected homology between the V21-like Orpheovirus sequence, Sputnik virophage V21 protein, the Zamilon protein, and a putative transferase present in the genomes of Mimivirus lineages A, B, and C. This V21-like protein also shared a common ancestor with all Sputnik virophages, and Zamilon virophage (**Figure 6**). No transposase or other mobile elements could be detected, no other special interest homology with other proteins was detected although a Ribonuclease III such as that in MIMIVIRE (Levasseur et al., 2016b) was present near this V21-like sequence in the genome of Orpheovirus.

DISCUSSION

Since the isolation of Faustovirus in 2015, all positive samples have been sewages samples or samples collected near to sewage areas (Reteno et al., 2015; Benamar et al., 2016; Bou Khalil et al., 2016; Cherif Louazani et al., 2017). We suspected that rats could also be a potential reservoir of Faustoviruses. In order to decipher the Faustovirus' reservoirs, and in attempt to study the viral frequency and persistence in the environment, notably during seasonality (Martínez et al., 2007; Johannessen et al., 2017), we decided to explore the same area of sampling 4 months later. We succeeded in re-isolating, in the same area, more Faustoviruses in sewage samples (data not shown) but not in rat stools samples, and a new giant virus was revealed, that we called Orpheovirus IHUMI-LCC2. This virus represents a new virus, the first to come with an ovoid form at a size higher than 1 μm isolated from *V. vermiformis* as a new host cell, and a genome of 1,473,573 bp largely exceeding the genomes of Cedratvirus A11, *P. massiliensis LC8,* and *P. sibericum*. Orpheovirus conserved a replicative cycle which is typical but delayed in terms of cell burst or complete lysis, which could be due to its host *V. vermiformis* showing different features regarding the routinely used host *Acanthamoeba* spp. (Andreani et al., 2017).

Although Orpheovirus appears to share some replicative elements and genomic bases with Cedratvirus A11, *P. massiliensis*

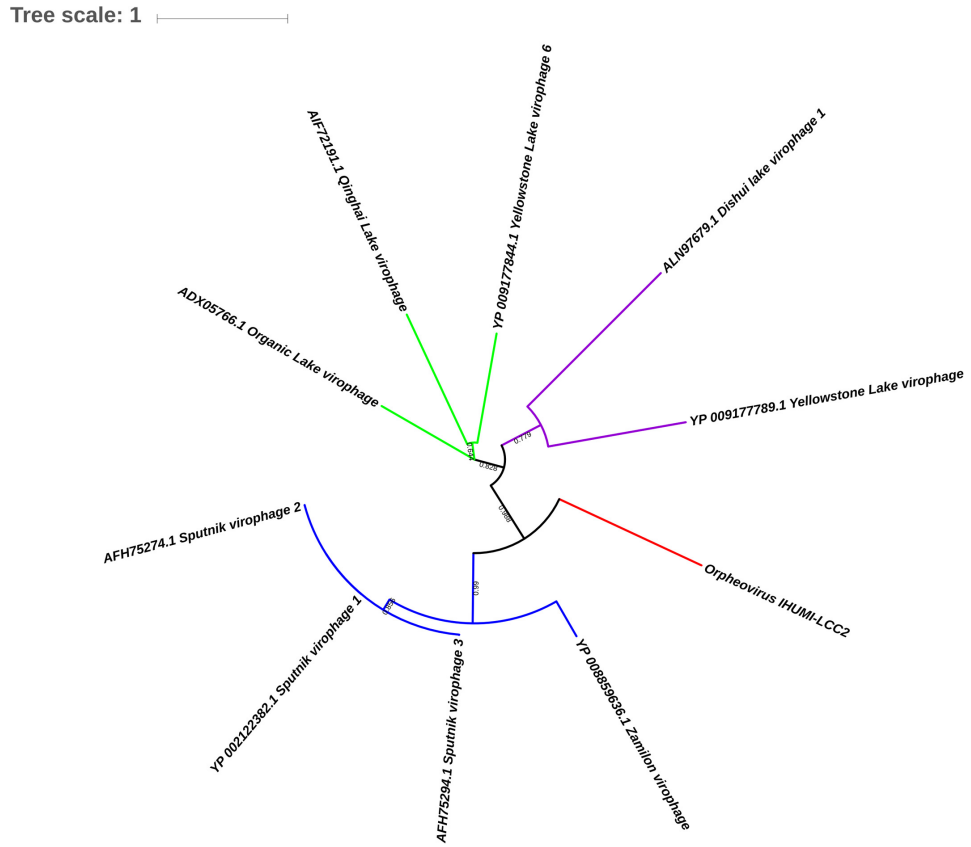


FIGURE 6 | Phylogenetic tree based on V21-like protein of virophages and Orpheovirus IHUMI-LCC2 sequences. 484 positions were used to build the tree. Branch values inferior to a bootstrap value of 0.5 were deleted.

LC8, and *P. sibericum*, some other elements highlighted a complete divergent evolution. With its genomic size, high number of paralogs, its eight aminoacyl tRNA synthetase aaRS, its low GC content, and its high number of ORFans (≈66% at a 10^{-5} *e*-value), we propose Orpheovirus as a potential member of a new putative family; the *Orpheoviridae* closely related to the recently proposed *Pithoviridae*. To do so, more viral descriptions including new isolates are needed to understand genomic links in these novel expanding and complex families. Nevertheless, some new viral descriptions such as that of Klosneuvirus (Schulz et al., 2017), have reported complete translational components, and this could create a broader understanding of the viral lifestyle and tRNA synthetase usages. In contrast, aminoacyl tRNA synthetases (aaRS) are frequently (but not entirely) found and described in isolated viruses (Abrahão et al., 2017), and we are still unaware of the viral benefits of possessing aminoacyl tRNA synthetase or tRNA during the infectious cycle. Simultaneously, and following the discovery of the MIMIVIRE system, it has become more widespread to search for virophages sequences in giant viruses genomes. We found a high conserved size and similar V21-like sequence in Orpheovirus that made us investigate the probability of an integrated virophage sequence in the Orpheovirus genome, as is the case for Mavirus (Fischer and Hackl, 2016) in its protist host *Cafeteria roenbergensis* or the Sputnik 2 virophage in the Lentille virus (Desnues et al., 2012). However, no mobile elements could be detected and no relationship could be found even when this sequence was closely located to the Ribonuclease III as is the case in MIMIVIRE. In contrast, the fact that the V21-like sequence of Orpheovirus, together with the V21 of Sputnik and Zamilon, showed homology with a putative transferase present in the genomes of Mimivirus lineages A, B, and C, led us to postulate that these sequences could either have a similar function to transferase or a protein that simply interacts with the putative transferase.

Despite all these findings, the description of Orpheovirus, along with the previous findings in Pandoraviruses, Pithoviruses, and Cedratviruses, has revealed a large range of viruses with various extraordinary ovoid shapes, which have expanded the research characteristics for viral isolation. Some more sewers should be investigated at different time stages or seasonal dates. In addition, animal stool samples should be more commonly considered as potential new reservoirs for giant viruses. Finally, a large part of this vast world of giant viruses is still unknown, particularly its evolution and ancestors. For this reason, more strains should be isolated and described, and more data is needed. It is likely that further descriptions will increase knowledge and diversity across the NCLDV.

AUTHOR CONTRIBUTIONS

JA and BL designed the study and experiments. JA, JK, EB, IH, CM, and AL performed the sample collection, virus isolation, experiments and/or analyses. JA, DR, and BL wrote the manuscript. All authors approved the final manuscript.

FUNDING

This work received a help of the ANR (National Agency for Research) through "future investments" program n°10-IAHU-03.

ACKNOWLEDGMENTS

The authors would particularly like to thank Aurélia Magnien for her help with the sample collection and Claire Andréani for her help in the improvement of the English correction.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2017.02643/full#supplementary-material

REFERENCES

Abergel, C., Legendre, M., and Claverie, J.-M. (2015). The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev.* 39, 779–796. doi: 10.1093/femsre/fuv037

Abrahão, J. S., Araújo, R., Colson, P., and La Scola, B. (2017). The analysis of translation-related gene set boosts debates around origin and evolution of mimiviruses. *PLOS Genet.* 13:e1006532. doi: 10.1371/journal.pgen.1006532

Aherfi, S., Colson, P., La Scola, B., and Raoult, D. (2016). Giant viruses of amoebas: an update. *Front. Microbiol.* 7:349. doi: 10.3389/fmicb.2016.00349

Andreani, J., Aherfi, S., Bou Khalil, J. Y., Di Pinto, F., Bitam, I., Raoult, D., et al. (2016). Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. *Viruses* 8:300. doi: 10.3390/v8110300

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

Antwerpen, M. H., Georgi, E., Zoeller, L., Woelfel, R., Stoecker, K., and Scheid, P. (2015). Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. *Genome Announc.* 3:e00136-15. doi: 10.1128/genomeA.00136-15

Arslan, D., Legendre, M., Seltzer, V., Abergel, C., and Claverie, J.-M. (2011). Distant mimivirus relative with a larger genome highlights the fundamental features of megaviridae. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17486–17491. doi: 10.1073/pnas.1110889108

Benamar, S., Reteno, D. G. I., Bandaly, V., Labas, N., Raoult, D., and La Scola, B. (2016). Faustoviruses: comparative genomics of new megavirales family members. *Front. Microbiol.* 7:3. doi: 10.3389/fmicb.2016.00003

Bertelli, C., Mueller, L., Thomas, V., Pillonel, T., Jacquier, N., and Greub, G. (2017). Cedratvirus lausannensis – digging into pithoviridaediversity. *Environ. Microbiol.* 19, 4022–4034. doi: 10.1111/1462-2920.13813

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29, 2607–2618. doi: 10.1093/nar/29.12.2607

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

Bou Khalil, J. Y., Andreani, J., Raoult, D., and La Scola, B. (2016). A Rapid strategy for the isolation of new faustoviruses from environmental samples using *Vermamoeba vermiformis*. *J. Vis. Exp.* 112:e54104. doi: 10.3791/54104

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21848–21853. doi: 10.1073/pnas.0911354106

Cherif Louazani, A., Andreani, J., Ouarhache, M., Aherfi, S., Baptiste, E., Levasseur, A., et al. (2017). Genome sequences of new Faustovirus strains ST1 and LC9, isolated from the South of France. *Genome Announc.* 5:e00613-17. doi: 10.1128/genomeA.00613-17

Colson, P., De Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D. K., et al. (2013). "Megavirales," a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* 158, 2517–2521. doi: 10.1007/s00705-013-1768-6

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

Contreras-Moreira, B., and Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* 79, 7696–7701. doi: 10.1128/AEM.02411-13

Da Cunha, V., Gaia, M., Gadelle, D., Nasir, A., and Forterre, P. (2017). Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. *PLOS Genet.* 13:e1006810. doi: 10.1371/journal.pgen.1006810

Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704

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

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

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

Filée, J. (2014). Multiple occurrences of giant virus core genes acquired by eukaryotic genomes: the visible part of the iceberg? *Virology* 466–467, 53–59. doi: 10.1016/j.virol.2014.06.004

Filée, J. (2015). Genomic comparison of closely related giant viruses supports an accordion-like model of evolution. *Front. Microbiol.* 6:593. doi: 10.3389/fmicb.2015.00593

Fischer, M. G., and Hackl, T. (2016). Host genome integration and giant virus-induced reactivation of the virophage mavirus. *Nature* 540, 288–291. doi: 10.1038/nature20593

Gaia, M., Benamar, S., Boughalmi, M., Pagnier, I., Croce, O., Colson, P., et al. (2014). Zamilon, a novel virophage with *Mimiviridae* host specificity. *PLOS ONE* 9:e94923. doi: 10.1371/journal.pone.0094923

Iyer, L. M., Aravind, L., and Koonin, E. V. (2001). Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.* 75, 11720–11734. doi: 10.1128/JVI.75.23.11720-11734.2001

Johannessen, T. V., Larsen, A., Bratbak, G., Pagarete, A., Edvardsen, B., Egge, E. D., et al. (2017). Seasonal dynamics of haptophytes and dsDNA algal viruses suggest complex virus-host relationship. *Viruses* 9:84. doi: 10.3390/v9040084

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Khalil, J. Y. B., Andreani, J., and La Scola, B. (2016a). Updating strategies for isolating and discovering giant viruses. *Curr. Opin. Microbiol.* 31, 80–87. doi: 10.1016/j.mib.2016.03.004

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

Khalil, J. Y. B., Robert, S., Reteno, D. G., Andreani, J., Raoult, D., and La Scola, B. (2016b). High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. *Front. Microbiol.* 7:26. doi: 10.3389/fmicb.2016.00026

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

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* 455, 100–104. doi: 10.1038/nature07218

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

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

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

Levasseur, A., Bekliz, M., Chabrière, E., Pontarotti, P., La Scola, B., and Raoult, D. (2016b). MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature* 531, 249–252. doi: 10.1038/nature17146

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44, W54–W57. doi: 10.1093/nar/gkw413

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 1:18. doi: 10.1186/2047-217X-1-18

Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331. doi: 10.1093/nar/gkh454

Martínez, J. M., Schroeder, D. C., Larsen, A., Bratbak, G., and Wilson, W. H. (2007). Molecular dynamics of *Emiliania huxleyi* and cooccurring viruses during two separate mesocosm studies. *Appl. Environ. Microbiol.* 73, 554–562. doi: 10.1128/AEM.00864-06

Moreira, D., and Brochier-Armanet, C. (2008). Giant viruses, giant chimeras: the multiple evolutionary histories of mimivirus genes. *BMC Evol. Biol.* 8:12. doi: 10.1186/1471-2148-8-12

Moreira, D., and López-García, P. (2015). Evolution of viruses and cells: do we need a fourth domain of life to explain the origin of eukaryotes? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140327. doi: 10.1098/rstb.2014.0327

Nadalin, F., Vezzi, F., and Policriti, A. (2012). GapFiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics* 13(Suppl. 14):S8. doi: 10.1186/1471-2105-13-S14-S8

Pagnier, I., Reteno, D.-G. I., Saadi, H., Boughalmi, M., Gaia, M., Slimani, M., et al. (2013). A decade of improvements in Mimiviridae and Marseilleviridae isolation from amoeba. *Intervirology* 56, 354–363. doi: 10.1159/000354556

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

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

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

Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., et al. (2017). Giant viruses with an expanded complement of translation system components. *Science* 356, 82–85. doi: 10.1126/science.aal4657

Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., et al. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179. doi: 10.1038/nature14447

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

Vitreschak, A. G. (2002). Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* 30, 3141–3151. doi: 10.1093/nar/gkf433

Yutin, N., Wolf, Y. I., and Koonin, E. V. (2014). Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* 466–467, 38–52. doi: 10.1016/j.virol.2014.06.032

Yutin, N., Wolf, Y. I., Raoult, D., and Koonin, E. V. (2009). Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol. J.* 6:223. doi: 10.1186/1743-422X-6-223

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358. doi: 10.1038/nature21031

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 April 2019
doi: 10.3389/fmicb.2019.00671

Microscopic Analysis of the *Tupanvirus* Cycle in *Vermamoeba vermiformis*

*Lorena C. F. Silva[1], Rodrigo Araújo Lima Rodrigues[1], Graziele Pereira Oliveira[1], Fabio Pio Dornas[2], Bernard La Scola[3], Erna G. Kroon[1] and Jônatas S. Abrahão[1]**

[1] Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [2] Faculdade de Ciências Básicas e da Saúde, Departamento de Farmácia, Universidade Federal do Vale do Jequitinhonha e Mucuri, Diamantina, Brazil, [3] Faculté de Médecine, Aix-Marseille Université, Marseille, France

Since *Acanthamoeba polyphaga mimivirus* (APMV) was identified in 2003, several other giant viruses of amoebae have been isolated, highlighting the uniqueness of this group. In this context, the tupanviruses were recently isolated from extreme environments in Brazil, presenting virions with an outstanding tailed structure and genomes containing the most complete set of translation genes of the virosphere. Unlike other giant viruses of amoebae, tupanviruses present a broad host range, being able to replicate not only in *Acanthamoeba* sp. but also in other amoebae, such as *Vermamoeba vermiformis*, a widespread, free-living organism. Although the *Tupanvirus* cycle in *A. castellanii* has been analyzed, there are no studies concerning the replication of tupanviruses in other host cells. Here, we present an in-depth microscopic study of the replication cycle of *Tupanvirus* in *V. vermiformis*. Our results reveal that *Tupanvirus* can enter *V. vermiformis* and generate new particles with similar morphology to when infecting *A. castellanii* cells. *Tupanvirus* establishes a well-delimited electron-dense viral factory in *V. vermiformis*, surrounded by lamellar structures, which appears different when compared with different *A. castellanii* cells. Moreover, viral morphogenesis occurs entirely in the host cytoplasm within the viral factory, from where complete particles, including the capsid and tail, are sprouted. Some of these particles have larger tails, which we named "supertupans." Finally, we observed the formation of defective particles, presenting abnormalities of the tail and/or capsid. Taken together, the data presented here contribute to a better understanding of the biology of tupanviruses in previously unexplored host cells.

Keywords: *Tupanvirus*, viral characterization, viral cycle, giant viruses, *Vermamoeba vermiformis*

OPEN ACCESS

Edited by:
Akio Adachi,
Kansai Medical University, Japan

Reviewed by:
Jonas Dutra Albarnaz,
University of Cambridge,
United Kingdom
Masaharu Takemura,
Tokyo University of Science, Japan

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

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

Received: *19 December 2018*
Accepted: *18 March 2019*
Published: *03 April 2019*

Citation:
Silva LCF, Rodrigues RAL, Oliveira GP, Dornas FP, La Scola B, Kroon EG and Abrahão JS (2019) Microscopic Analysis of the Tupanvirus Cycle in Vermamoeba vermiformis. Front. Microbiol. 10:671. doi: 10.3389/fmicb.2019.00671

INTRODUCTION

Since the isolation of *Acanthamoeba polyphaga mimivirus* (APMV) in the early 2000s, giant viruses have been arousing interest due to their structural, biological, and genomic complexity (La Scola et al., 2003; Colson et al., 2017). Since then, questions have been raised about the relationship of these viruses to their hosts, their evolution, and their position in the microbial world. After about 15 years of study, several other giant viruses of amoebae were isolated, such as the marseilleviruses, pandoraviruses, and pithoviruses, among others, contributing further knowledge about the diversity of this group (Colson et al., 2017). Many other interesting and unusual viruses can be spread across a wide range of environments, so the discovery and characterization of these viruses is still a promising field and a major challenge (Colson et al., 2017).

In 2015, the prospection of giant viruses from 17 samples from soda lakes and oceanic soil sediments collected in Brazil was performed, resulting in the isolation of two new viral isolates, named *Tupanvirus* soda lake (TPVsl) and *Tupanvirus* deep ocean, which are able to replicate in amoebae of different genera, such as *Acanthamoeba* and *Vermamoeba*, among others (Abrahão et al., 2018). Due to their genetic and phylogenetic characteristics, tupanviruses are proposed to be members of the family Mimiviridae, constituting a new genus "*Tupanvirus*" (Abrahão et al., 2018; Rodrigues et al., 2018). The biological characterization of *Tupanvirus* strains showed a peculiar structure. A capsid similar to that of a *Mimivirus* with the stargate portal on one side and surrounded by fibrils (Zauberman et al., 2008; Abrahão et al., 2018). However, the presence of a cylindrical tail attached to the capsid in the isolates, which can extend their sizes to more than 2 μm, seems to be the distinguishing feature of *Tupanvirus* particles compared with other giant viruses described until now (Abrahão et al., 2018). Mimiviruses attracted attention due to the presence of a large, icosahedral capsid associated with fibrils; pandoraviruses, cedratviruses, and pithoviruses show an ovoid morphology, are very large viruses, and have apical pores; however, in none of these viruses was there any structure resembling that of a tail, which is only found in tupanviruses (La Scola et al., 2003; Philippe et al., 2013; Legendre et al., 2014; Abrahão et al., 2018).

To date, the replication cycle of a *Tupanvirus* strain, TPVsl, has been analyzed in *A. castellanii* by electron microscopy, among other techniques (Abrahão et al., 2018). The analyses showed that the particles bind to the surface of the amoeba and penetrate the cell, likely by a phagocytic process. The stargate opens, and the inner capsid and tail membranes merge with the phagosomal membrane, releasing the genome into the cell cytoplasm. A viral factory of the volcano type is formed, wherein the genome replication and morphogenesis of new particles occur, as described for other mimiviruses (Suzan-Monti et al., 2007; Abrahão et al., 2018). The tail of the particle is supposedly attached to the capsid after its formation and closure, although there is no clear evidence about this step of *Tupanvirus* morphogenesis (Abrahão et al., 2018). In late stages of the cycle, the amoebic cytoplasm is filled with several viral particles, followed by cell lysis and particle release (Abrahão et al., 2018).

As tupanviruses were the first giant amoeba viruses that demonstrated this ability to replicate in protozoa belonging to different genera, this study aimed to analyze in detail the replication cycle of TPVsl in *V. vermiformis* to elucidate and compare the steps of its replication cycle with those already evidenced in *A. castellanii* and other aspects that still remain unclear.

MATERIALS AND METHODS

Virus Preparation and Cells

Tupanvirus soda lake (TPVsl) was isolated from a soda lake sample from the Pantanal region in Brazil and was produced and purified as previously described (Abrahão et al., 2018). Briefly, *A. castellanii* (ATCC 30010) cells were grown in 75 cm^2 cell culture flasks (Nunc, United States) in peptone–yeast extract–glucose (PYG) medium (Visvesvara and Balamuth, 1975) supplemented with 25 mg/mL fungizone (Amphotericin B, Cristalia, Brazil), 500 U/mL penicillin, and 50 mg/mL gentamicin (Schering-Plough, Brazil). After reaching confluence, the amoebae were infected at a multiplicity of infection (m.o.i) of 0.1 and incubated at 32°C until cytopathic effects (CPE) were observed. Supernatants from the infected amoebae were collected and filtered through a 0.8 μm filter to remove cell debris. The viruses were purified by centrifugation through a sucrose cushion (22%), suspended in phosphate-buffered saline (PBS), and stored at −80°C.

Asynchronous Cycle of TPVsl in *V. vermiformis* and Transmission Electron Microscopy (TEM)

To investigate the asynchronous cycle of TPVsl in *V. vermiformis* cells (ATCC 20237), 25 cm^2 cell culture flasks with 5×10^6 of *V. vermiformis* in 10 mL of PYG medium were infected with TPVsl at a m.o.i. of 0.1 and incubated at 32°C for 36 h. After the period of infection, the cells were collected and submitted to three cycles of freezing (−80°C)/ thawing (25°C) for cell lysis and virus release. Samples were then clarified for total particles counting in Neubauer Chamber and for titration. The viral titer was determined using the $TCID_{50}$ (tissue culture infective dose) method that was calculated using the Reed and Muench (1938) method in 96-well plates with 4×10^4 amoebae per well. The rest of the cells were prepared for microscopy assays. For this, the infected *V. vermiformis* cells were collected, pelleted by centrifugation at 1500 *g* for 10 min, and fixed in microcentrifuge tubes with 1 mL of 2.5% glutaraldehyde solution in 0.1 M sodium phosphate buffer pH 7,4 for 1 h at room temperature. The samples were then washed three times with 0.1 M sodium phosphate buffer, post-fixed with 2% osmium tetroxide, and embedded in Epon resin. Ultrathin sections were then analyzed under TEM (Spirit BioTWIN FEI, 120 kV) at the Center of Microscopy of UFMG.

Scanning Electron Microscopy

For analysis under scanning electron microscopy (SEM), the infected *V. vermiformis* cells were collected after 24–36 h of infection, lysed by freezing/thawing and pelleted by centrifugation at 1500 *g* for 10 min. After, they were added to round glass coverslips covered with poly-L-lysine, and fixed with 2.5% glutaraldehyde solution in 0.1 M cacodylate buffer pH 7,4 for 1 h at room temperature. The samples were then washed three times with 0.1 M cacodylate buffer and post-fixed with 1.0% osmium tetroxide for 1 h at room temperature. After a second fixation, the samples were washed three times with 0.1 M cacodylate buffer and immersed in 0.1% tannic acid for 20 min. The samples were then washed in cacodylate buffer and dehydrated by serial passage in ethanol solutions at concentrations ranging from 35 to 100%. Samples were subjected to critical point drying using CO_2, placed in stubs and metallized with a 5-nm gold layer. The analyses were completed using SEM (FEG Quanta 200 FEI) at the Center of Microscopy of UFMG.

TPVsl One-Step-Growth-Curve in *V. vermiformis*

To get one-step-growth-curve of TPV in *V. vermiformis*, 25-cm^2 cell culture flasks with 5×10^6 cells of *V. vermiformis* in 10 mL of PYG medium were infected with TPVsl at M.O.I. of 10 and incubated at 32°C. At different time points, the flasks were observed by light microscope to monitor the evolution of the CPE. Moreover, cells were collected and used for titration by $TCID_{50}$ as described above.

RESULTS

The Early Steps of *Tupanvirus* Infection in *Vermamoeba vermiformis*

To evaluate the replication profile of TPVs1 in *V. vermiformis*, asynchronous infections were performed, and the infected cells were prepared for electron microscopy analyses. Our first images revealed that the TPV particles in *Vermamoeba* cells acquire the same structure that was observed in *Acanthamoeba* cells (**Figure 1**) as described by Abrahão et al. (2018). Tupanviruses present a capsid of about 450 nm, similar to that of mimiviruses, including a stargate region in one of the vertexes and multiple layers, including an electron-dense structure inside the capsid, indicative of a lipid membrane. Attached to that capsid basis, there is a cylindrical tail approximately 450 nm in diameter and 550 nm in length, which increase the size of the virus, a unique feature displayed by these viruses (**Figure 1**). The complete viruses are approximately 1.2 μm, although some particles can be longer, reaching over 2.0 μm due to variation in tail size. We termed these larger particles "supertupans" (**Figure 1C**). Curiously, these enormous particles were observed recurrently in both our TEM and SEM preparations of infected *V. vermiformis* cells.

Regarding the initial steps of the *Tupanvirus* cycle in *V. vermiformis*, the particles attach themselves to the surface of the host cell and penetrate the cell, most likely through phagocytosis, as amoebic pseudopods encompassing viruses close to their surface are observed (**Figures 2A,B**). Inside the amoeba cytoplasm, viruses stay within phagosomes, normally with one particle per phagosome, although multiple viruses can enter the host cell simultaneously, resulting in more than one particle inside a phagosome (**Figures 2C–E**). In these initial steps, we could observe the amoeba nucleus clearly without apparent changes, with the nucleolus highly evident (**Figure 2D**). Furthermore, mitochondria and several vacuoles were observed around the internalized viruses (**Figure 2E**). After, the stargate opened, and the viral capsid inner membrane fused with the phagosomes membrane, culminating with the release of the genome into the host cytoplasm (**Figure 2F**).

Analysis of the *Tupanvirus* Viral Factory in *Vermamoeba vermiformis*

Giant viruses usually establish delimited regions in the hosts' cytoplasm, named viral factories (VF) (Kuznetsov et al., 2013; Mutsafi et al., 2014; Andrade et al., 2017). Tupanviruses are no exception, as previously demonstrated upon infection of *A. castellanii* cells (Abrahão et al., 2018). According to our observations, TPVsl also establish a VF in the host cytoplasm when infecting *V. vermiformis* cells. After the early steps of infection, we observed the generation of a VF, a structure well delimited in the host's cell cytoplasm (**Figure 3A**). The VF formed upon infection of *V. vermiformis* has a peculiar appearance (**Figures 3A,B**). Its margin is more delimited and irregular than

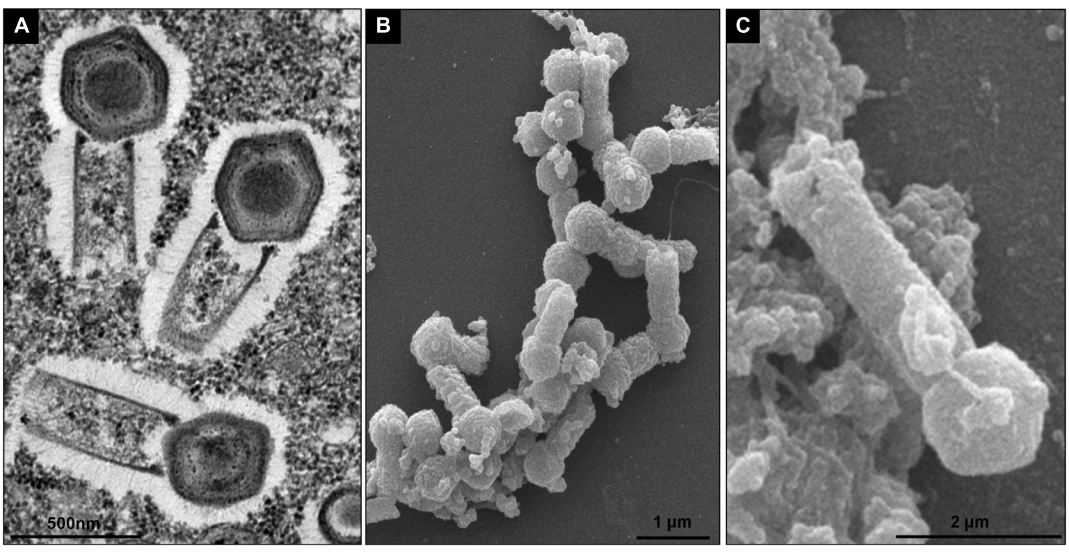


FIGURE 1 | *Tupanvirus* soda lake particle. **(A)** Mature particle of Tupanvirus soda lake (TPVsl) in *V. vermiformis* under transmission electron microscopy (TEM). **(B)** Mature particle of TPVsl in *V. vermiformis* under scanning microscopy. Is it possible note the peculiar TPVsl morphology, with the tail attached to a *Mimivirus*-like capsid. **(C)** "Supertupan" in *V. vermiformis* under scanning electron microscopy.

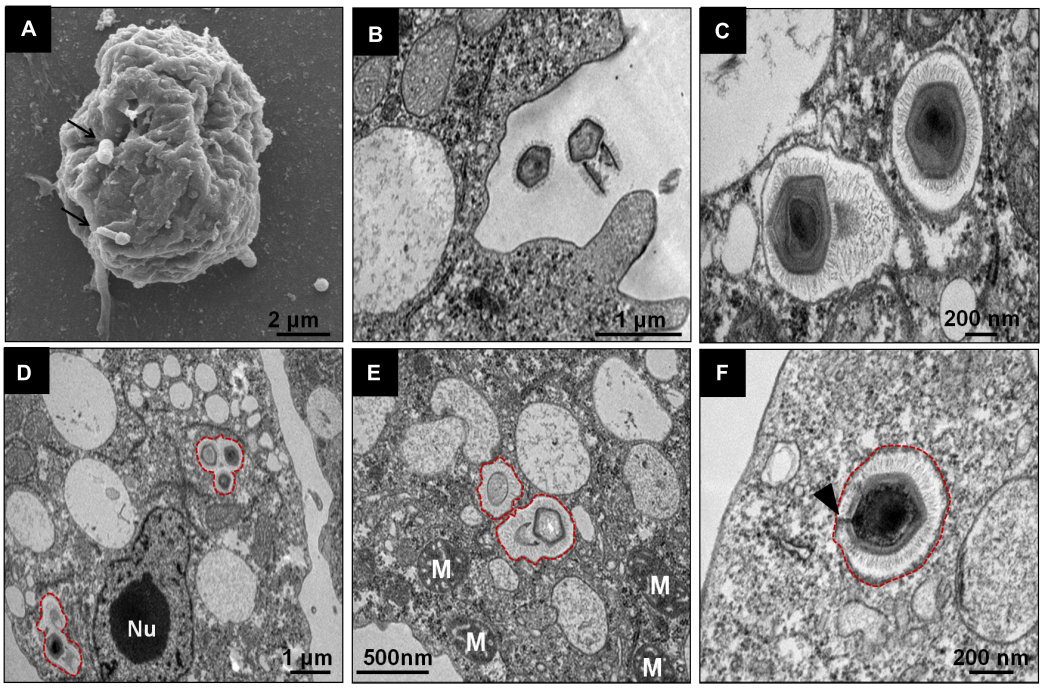


FIGURE 2 | Initial steps of the replication cycle of *Tupanvirus* soda lake in *V. vermiformis*. **(A)** Detail of TPVsl particles attached to a *V. vermiformis* cell. **(B)** Amoebae emit pseudopodia to encompass viral particles that are internalized through phagocytosis. **(C–E)** Details of viral particles that remain within phagosomes (red outlines). The cell nucleus remains apparent/electrodense, and the cytoplasm presents several empty vacuoles. **(F)** During the uncoating step, the stargate opens, followed by membrane fusion. The viral capsid, indicated by the black arrow, releases the genome. M, mitochondria; Nu, nucleus.

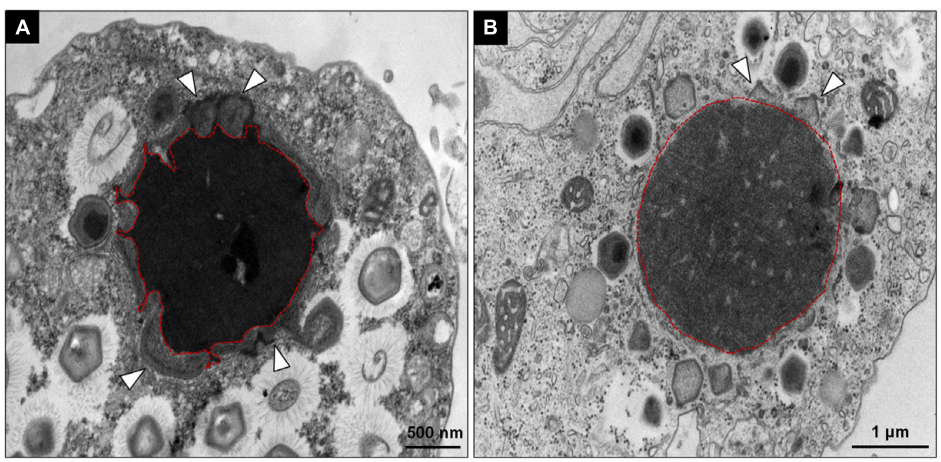


FIGURE 3 | Comparison between the TPVsl viral factory (VF) in *V. vermiformis* and *A. castellanii*. **(A)** Mature VF in *V. vermiformis*. **(B)** Mature VF in *A. castellanii*. It is possible to note that the VF in the *Vermamoeba* on the left is electron-dense and has an irregular border in the periphery (red), from which sprouts the structures to form new viral particles. The budding of viral capsids is highlighted by the white arrows.

in *A. castellanii*, evidencing a lamellar aspect of the VF in its mature stage (**Figure 3A**). The structure seems to be formed by several layers that expand in an way analogous to that of crescents described for poxviruses and marseilleviruses from where the viral structures sprout (Maruri-Avidal et al., 2011; Andrade et al., 2017). It is worthy of note that the VF of giant viruses is the region wherein the viral genome is replicated, and new particles are assembled (Kuznetsov et al., 2013; Mutsafi et al., 2014; Andrade et al., 2017). For that reason, several particles are expected to be found in these regions. This is valid for tupanviruses, since dozens of particles were observed in different SEM images (**Figure 4**). The particles appeared to be partially assembled, composed of a capsid and tail (**Figure 4A**), but fibrils were probably absent, since the stargate structure could be observed easily protruding

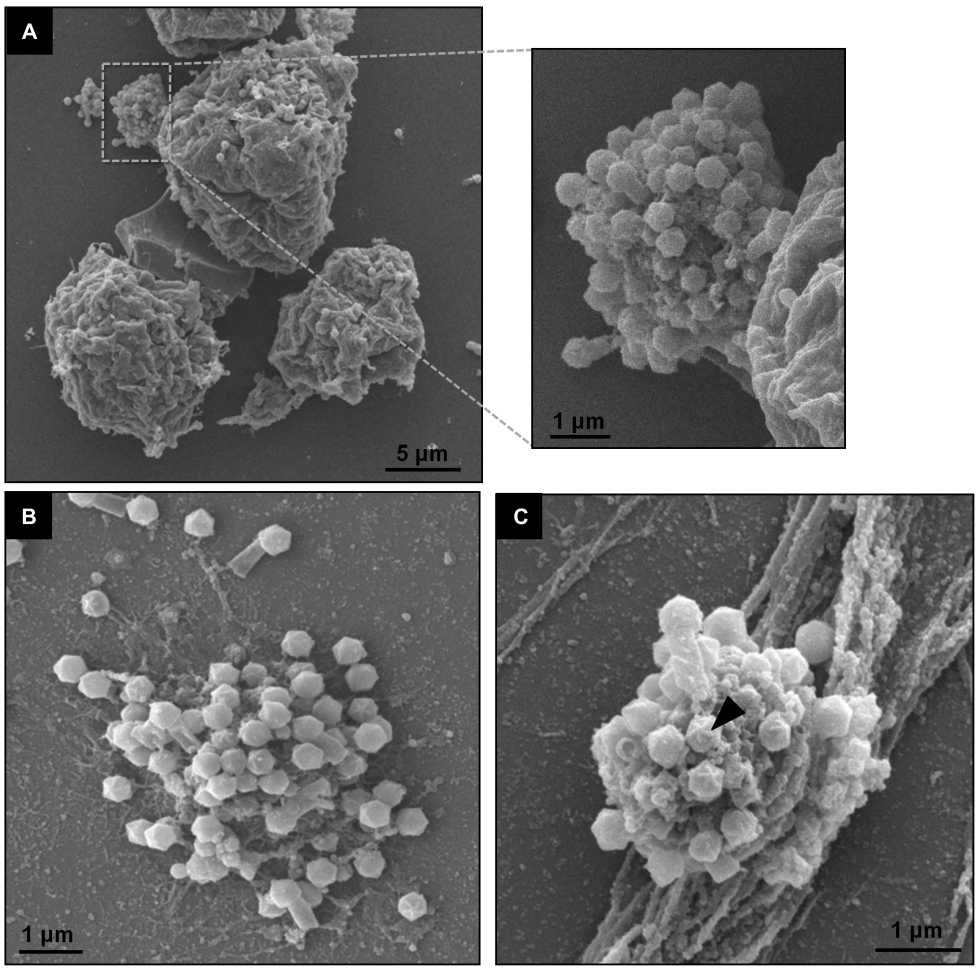


FIGURE 4 | TPVsl viral factory by scanning microscopy. **(A)** Details of a mature VF involved in viral morphogenesis, including tail attachment. **(B,C)** Isolated VFs releasing viral particles. The stargate in a capsid is indicated by the black arrow.

in many particles, indicating that the particles, and the VF undergoes different levels of maturation (**Figures 4B,C**). At later stages of viral infection, once the VF is fully established, viral capsids are assembled, and the genome is incorporated at the periphery of the VF (**Figures 5A–C**). This event can occur before or after fibril acquisition, thus events are not likely to occur in chronological order (**Figures 5D,E**). In contrast to mimiviruses, it is likely that no particular area for fibril acquisition is formed during *Tupanvirus* VF maturation (Andrade et al., 2017). The viral tails apparently attach to the capsid immediately after genome incorporation and sprout from the VF along with the capsid, forming complete virions (**Figure 5F**).

The Final Step of the TPVsl Cycle in *V. vermiformis* Is Associated With Defective Particle Release

During the final step of the replication cycle, we observed a large increase in the number of typical TPVsl particles filling the cytoplasm, i.e., particles presenting a tailed capsid covered by fibrils and a size of approximately 1.2 µm (**Figures 6A–C**). Viral progeny formed by mature and complete particles accumulated in the amoebae cytoplasm and their release was mediated by cell lysis (**Figure 6D**). However, our analysis showed that this step is also associated with a high proportion of defective particles in *V. vermiformis* cells. Many images have shown that in some amoebae, the VF in its final stage presents a differentiated aspect: it is smaller, becomes less electron-dense, and loses its lamellar aspect (**Figures 7A–C**). This seems to be closely related to the budding of abnormal structures forming abnormal particles in the cytoplasm. In our analysis, we observed defective capsids without the expected pseudo-icosahedral symmetry and also not completely closed or surrounded by fibrils (**Figures 7A–E**). Furthermore, at this step we also observed defectives supertupans. Long tails are commonly noticed, and sometimes the cylindrical shape is replaced by undefined forms (**Figure 7F**). This process occurs in amoebae with final-stage mature VF. The comparative analysis of total particles and titrated particles obtained at the end of the asynchronous cycle showed that the number of total

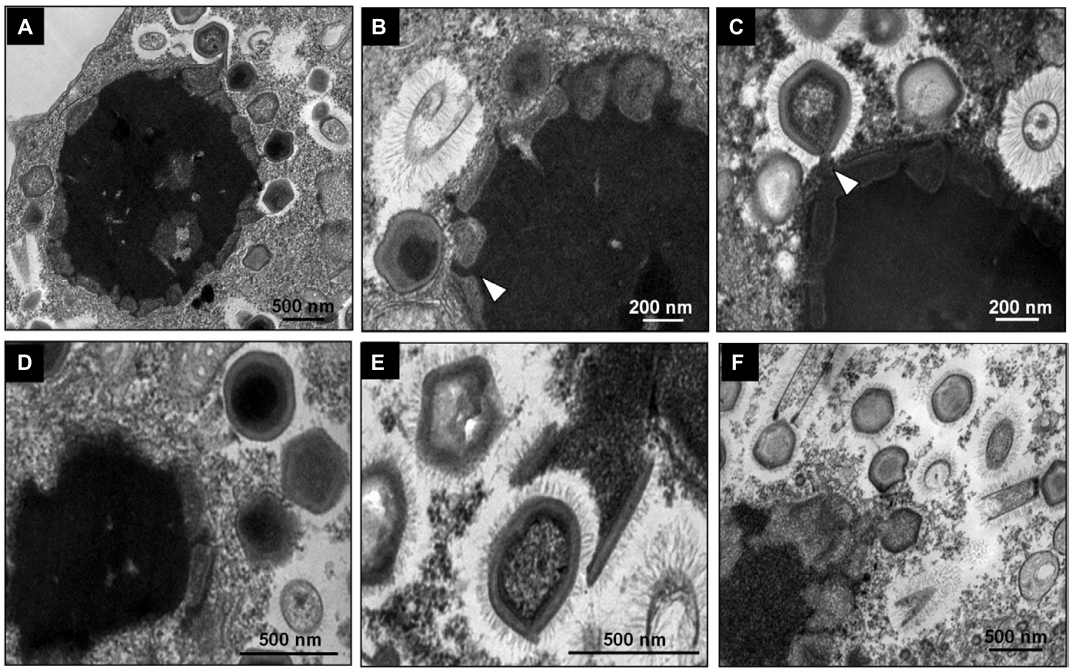


FIGURE 5 | Morphogenesis of *Tupanvirus* particles. **(A)** Overview of a mature VF and the budding of the structures that compose the viral progeny through its periphery. **(B)** Genome acquisition (shown by the arrows) by a capsid without fibrils. **(C)** Genome acquisition by a capsid with fibrils. **(D–F)** Details of the sprouting of viral capsids and tails through the VF.

particles is about two times higher in relation to the infectious particles (**Figure 7H**).

Characterization of CPE and Evolution of Viral Titer During Synchronous Infection

In order to characterize CPE triggered by *Tupanvirus* in *V. vermiformis*, that cells were infected at an m.o.i. of 10 and observed at up to 72 h.p.i. We observed that the formation of the CPE seems be slower in *V. vermiformis* than to that previously observed in *A. castellanii* (Abrahão et al., 2018). We observed that TPV induces in *V. vermiformis* cell rounding and early cluster formation, the typical "bunches" formed by TPV in amoeba (Oliveira et al., 2019), being visible only around 12 h.p.i., being most evident at 16 and 24 h.p.i. At 36–72 h p.i., we observed bunches disaggregation and lysis (**Supplementary Figure S1A**). One-step-growth-curve analysis revealed eclipse phase around 4 h.p.i. At 36 h.p.i., TPV title increases approximately 1 log (**Supplementary Figure S1B**) if compared to eclipse phase (4 h.p.i.), we observed titer increased about 3 log at 36 h.p.i.

DISCUSSION

Tupanviruses were isolated from extreme environments in Brazil and showed unprecedented characteristics, including the ability to replicate in different genera of protozoa (Abrahão et al., 2018). Our data suggest that TPV cycle in *V. vermiformis* is slower and less productive than TPV replication in *A. castellanii* (Oliveira et al., 2019; **Supplementary Figure S1**). The reason why we observe a delay in the evolution of TPV CPE in *V. vermiformis* requires more investigation as well. A similar profile was observed in the early phase of the replication cycle of TPVsl in *V. vermiformis* in relation to that which occurs in *A. castellanii*, with viral attachment to the amoeba surface and entry through phagocytosis (Abrahão et al., 2018). It is possible that tupanviruses attach to host cells by interaction of their fibrils with different glycans present on the cell surface, in a similar way to that observed for mimiviruses, although its composition remains to be elucidated (Rodrigues et al., 2015). The strategy of penetration by phagocytosis has recurrently been assumed for different giant viruses of amoebae, considering the size of the viral particles (larger than 500 nm) and the phagotrophic nature of amoebae (Suzan-Monti et al., 2007; Abrahão et al., 2014). However, it has been suggested that particles from smaller amoebae viruses such as marseilleviruses (approximately 250 nm) would not use this strategy but would use the other endocytic pathway or penetrate through phagocytosis when forming vesicles containing a large number of viral particles (Arantes et al., 2016). The phagocytic strategy for penetration was biologically demonstrated for APMV and *Cedratvirus getuliensis* by the use of pharmacological inhibitors of the phagocytosis process, demonstrating a considerable decrease in viral particle incorporation and replication success (Andrade et al., 2017; Silva et al., 2018). By observation of several TEM images, we suggest that this same strategy is adopted by TPVsl (**Figures 2A,B**), although other mechanisms, such as macropinocytosis, cannot be discarded for the moment.

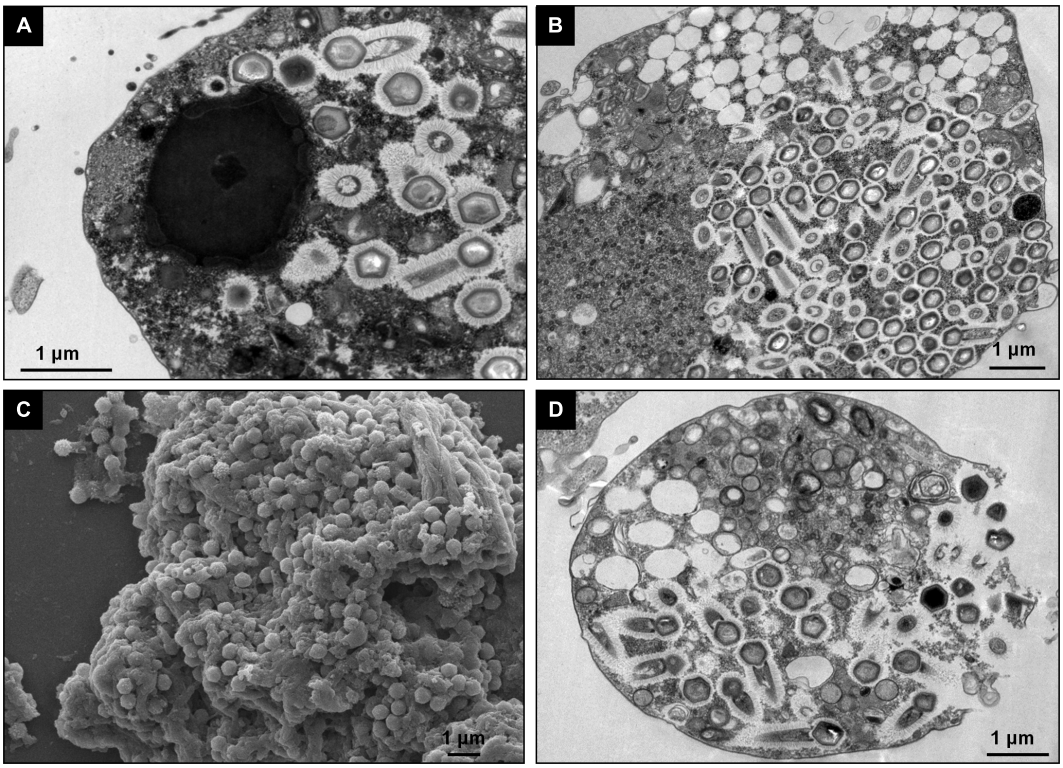


FIGURE 6 | Final steps of the replication cycle of *Tupanvirus* in *V. vermiformis*. **(A)** VF in the mature stage releasing mature viral particles shown by TEM. **(B,C)** Amoebae filled with mature viral particles shown by transmission **(B)** and scanning **(C)** electron microscopy. **(D)** Amoeba cell filled with new viral particles under lysis shown by TEM.

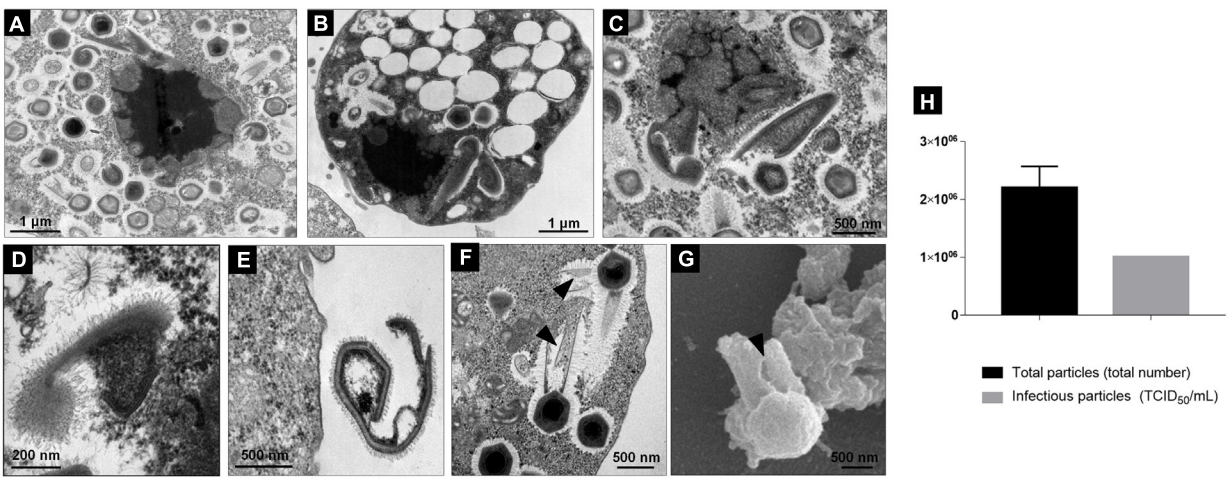


FIGURE 7 | Defective particles of *Tupanvirus* formed during its cycle in *V. vermiformis*. **(A–C)** VF in the last step of maturation releasing mature and defective particles shown by TEM. **(D,E)** Details of defective particles shown by TEM. **(F)** Details of "supertupans" with defective tails (indicated by the black arrow) shown by TEM. **(G)** A defective tail of TPVsl show by scanning microscopy. **(H)** Proportion of total particles and infectious particles during the asynchronous cycle of TPVsl in *V. vermiformis*.

The replication cycle appears to be entirely cytoplasmic, with the establishment of a well-defined VF, as previously reported for other related large DNA viruses (Mutsafi et al., 2010, 2014; Kuznetsov et al., 2013; Andrade et al., 2017). On the other hand, pandoraviruses are amoeba viruses that have a replication cycle involving the host nucleus in some way, due to the lack of genes essential for DNA replication in its genome, even though a large VF is observed (Philippe et al., 2013;

Andrade et al., 2018). In this context, we noticed a difference in the aspect of the VF on the two amoeba cells infected by TPVsl (**Figure 3**). However, this characteristic should be observed with caution. This may be due to some particular property of this amoeba, including how it reacts to TEM preparation. And also, because asynchronous cycle was used, it is possible that the differences in the VF reflect different stages of the viral morphogenesis. Several studies with other amoeba giant viruses have demonstrated this close relationship between VF and morphogenesis (Suzan-Monti et al., 2007; Mutsafi et al., 2010, 2014; Kuznetsov et al., 2013; Andrade et al., 2017). For APMV, it was demonstrated that the assembly of capsids from increasing lamellar structures starts in the periphery of the VF, followed by membrane biogenesis and then genetic material packing on the opposite side of the stargate and simultaneous fibril acquisition by passage through a less electron-dense area surrounding the VF (Mutsafi et al., 2014; Andrade et al., 2017). For tupanviruses, we observed that the assembled capsid containing its various layers can be filled with DNA before or after fibril acquisition, since the VF of TPVsl does not present a delimited area for this event, neither in *V. vermiformis* nor in *A. castellanii*, in contrast to the observed for mimiviruses (**Figure 5**; Andrade et al., 2017).

Viral morphogenesis is a complex process during the replication cycle of a virus, in particular for the large DNA viruses, which involves the presence of many different and large structures (Moss, 2013; Andrade et al., 2017; Silva et al., 2018). Furthermore, some DNA viruses, such as herpesviruses, poxviruses, and mimiviruses, incorporate transcripts into their forming particles during this step (Raoult et al., 2004; Grossegesse et al., 2017). Recently, a next-generation sequencing (NGS) study showed that the content of transcripts incorporated by cowpox virus intracellular mature virion (IMV) in human cells (Hep-2) or murine cells (Rat-2), is not identical and thus may be due to host-specific incorporation (Grossegesse et al., 2017). Although no significant differences could be observed in the TPV cycle in *Acanthamoeba* and *Vermamoeba* concerning viral morphogenesis, we still cannot affirm that the content of transcripts and proteins in virions during their formation is the same in both cells. Further comparative studies involving genomics and proteomics would bring forward valuable information on this subject.

Viral progeny release is mediated by cell lysis in a similar way as previously demonstrated for other giant viruses (**Figure 6**; Abrahão et al., 2014). An interesting fact that drew attention at this step was the greatest presence of defective particles in *V. vermiformis* cytoplasm (**Figure 7**). This has already been verified for APMV in *A. castellanii* cells, suggesting that defective particles in giant viruses are not only formed in the presence of virophages but can also be an event associated with the normal replication cycle (La Scola et al., 2008; Andrade et al., 2017). Also, our data demonstrate that the proportion of total particles is about two-fold higher than the number of infectious particles after an asynchronous cycle, highlighting the presence of defective particles and corroborates with the observed images (**Figure 7**). The reason why there appears to be more defective TPV particles following infection in *Vermamoeba* requires further investigation. Considering that tupanviruses have a broad spectrum of hosts, in contrast to other giant amoeba viruses, it is possible that the level of adaptability of the viruses in different amoeba genera or species can influence this profile. In conclusion, the data presented here contribute to a better understanding of the biology of tupanviruses in *V. vermiformis*.

AUTHOR CONTRIBUTIONS

LS, RR, FD, and GO performed experiments. JA, BLS, and EK designed the study. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We would like to thank colleagues from Gepvig, Laboratório de Vírus, IHU-Aix Marseille University and Microscopy Center of UFMG for their excellent support. We also would like to thank CAPES, FAPEMIG, MS, and CNPq for financial support (Decit/SCTIE/MoH). EK, BLS, JA, RR, and LS are members of a CAPES-COFECUB project. JA and EK are CNPq researchers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00671/full#supplementary-material

FIGURE S1 | Evolution of TPVsl cytopathic effect and infectious particles during the synchronous cycle. **(A)** *V. vermiformis* was infected with TPVsl using a high m.o.i. and visualized by light microscopy. We observed the formation of bunches after 12 h.p.i., that were disaggregated about 36 h.p.i. After this time, we observed lysis, but it was not total. The flasks were observed using the 100× objective on a light microscopy. **(B)** TPVsl one-step growth curve in *V. vermiformis* at an m.o.i. of 10. Error bars indicate standard deviation.

REFERENCES

Abrahão, J. S., Dornas, F. P., Silva, L. C. F., Almeida, G. M., Boratto, P. V. M., Colson, P., et al. (2014). Acanthamoeba polyphaga mimivirus and other giant viruses: an open field to outstanding discoveries. *Virol. J.* 11:120. doi: 10.1186/1743-422X-11-120

Abrahão, J. S., Silva, L. C. F., Silva, L. K. S., Khalil, J. Y. B., Rodrigues, R., Arantes, T., et al. (2018). Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* 9:749. doi: 10.1038/s41467-018-03168-1

Andrade, A. C. S. P., Boratto, P. V. M., Rodrigues, R. A. L., Bastos, T., Azevedo, B. L., Dornas, F. P., et al. (2018). New isolates of pandoraviruses: contribuition to the study of replication cycle steps. *J. Virol.* 93:e1942-18.

Andrade, A. C. S. P., Rodrigues, R. A. L., Oliveira, G. P., Andrade, K. R., Bonjardim, C. A., La Scola, B., et al. (2017). Filling knowledge gaps for mimivirus entry, uncoating, and morphogenesis. *J. Virol.* 91, e1335-17. doi: 10.1128/JVI.01335-17

Arantes, T. S., Rodrigues, R. A. L., Dos Santos Silva, L. K., Oliveira, G. P., De Souza, H. L., Khalil, J. Y., et al. (2016). The large marseillevirus explores different entry pathways by forming giant infectious vesicles. *J. Virol.* 11, 5246–5255. doi: 10.1128/JVI.00177-16

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

Grossegesse, M., Doellinger, J., Haldemann, B., Schaade, L., and Nitsche, A. (2017). A next-generation sequencing approach uncovers viral transcripts incorporated in poxvirus virions. *Viruses* 9, 296. doi: 10.3390/v9100296

Kuznetsov, Y. G., Klose, T., Rossmann, M., and Mcpherson, A. (2013). Morphogenesis of mimivirus and its viral factories: an atomic force microscopy study of infected cells. *J. Virol.* 20, 11200–11213. doi: 10.1128/JVI.01372-13

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

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* 7209, 100–104. doi: 10.1038/nature07218

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral Dna viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Maruri-Avidal, L., Domi, A., Weisberg, A. S., and Moss, B. (2011). Participation of vaccinia virus L2 protein in the formation of crescent membranes and immature virions. *J. Virol.* 85, 2504–2511. doi: 10.1128/JVI.02505-10

Moss, B. (2013). Poxvirus DNA replication. *Cold Spring Harb. Perspect. Biol.* 5:a010199. doi: 10.1101/cshperspect.a010199

Mutsafi, Y., Fridmann-Sirkis, Y., Milrot, E., Hevroni, L., and Minsky, A. (2014). Infection cycles of large DNA viruses: emerging themes and underlying questions. *Virology* 466–467, 3–14. doi: 10.1016/j.virol.2014.05.037

Mutsafi, Y., Zauberman, N., Sabanay, I., and Minsky, A. (2010). Vaccinia-like cytoplasmic replication of the giant Mimivirus. *Proc. Natl. Acad. Sci. U.S.A.* 13, 5978–5982. doi: 10.1073/pnas.0912737107

Oliveira, G. P., Silva, L. C. F., Leão, T., Mougari, S., Da Fonseca, F. G., Kroon, E. G., et al. (2019). Tupanvirus-infected amoebas are induced to aggregate with uninfected cells promoting viral dissemination. *Sci. Rep.* 9:183. doi: 10.1038/s41598-018-36552-4

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

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004). The 1.2-megabase genome sequence of mimivirus. *Science* 306, 1344–1350. doi: 10.1126/science.1101485

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty per cent end points. *Am. J. Epidemiol.* 27, 493–497. doi: 10.1093/oxfordjournals.aje.a118408

Rodrigues, R. A. L., Dos Santos Silva, L. K., Dornas, F. P., De Oliveira, D. B., Magalhães, T. F., Santos, D. A., et al. (2015). mimivirus fibrils are important for viral attachment to the microbial world by a diverse glycoside interaction repertoire. *J. Virol.* 89, 11812–11819. doi: 10.1128/JVI.01976-15

Rodrigues, R. A. L., Mougari, S., Colson, P., La Scola, B., and Abrahão, J. S. (2018). "Tupanvirus", a new genus in the family Mimiviridae. *Arch. Virol.* 164, 325–331. doi: 10.1007/s00705-018-4067-4

Silva, L. K. S., Andrade, A. C. S. P., Dornas, F. P., Rodrigues, R. A. L., Arantes, T. S., Kroon, E. G., et al. (2018). Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. *Sci. Rep.* 8:4000. doi: 10.1038/s41598-018-22398-3

Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L., and Raoult, D. (2007). Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga mimivirus. *PLoS One* 2:e328. doi: 10.1371/journal.pone.0000328

Visvesvara, G. S., and Balamuth, W. (1975). Comparative studies on related free-living and pathogenic amoebae with special reference to Acanthamoeba. *J. Protozool.* 22, 245–256. doi: 10.1111/j.1550-7408.1975.tb05860.x

Zauberman, N., Mutsafi, Y., Halevy, D. B., Shimoni, E., Klein, E., Xiao, C., et al. (2008). Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus. *PLoS Biol.* 6:e114. doi: 10.1371/journal.pbio.0060114

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 March 2018
doi: 10.3389/fmicb.2018.00463

Experimental Inoculation in Rats and Mice by the Giant Marseillevirus Leads to Long-Term Detection of Virus

Sarah Aherfi[1], *Claude Nappez*[1], *Hubert Lepidi*[1,2], *Marielle Bedotto*[1], *Lina Barassi*[1], *Priscilla Jardot*[1], *Philippe Colson*[1], *Bernard La Scola*[1], *Didier Raoult*[1] *and Fabienne Bregeon*[1,3]*

[1] *Institut Hospitalo Universitaire Méditerranée Infection, Assistance Publique-Hôpitaux de Marseille, Centre Hospitalo Universitaire Timone, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Marseille, France,* [2] *Laboratoire d'Anatomopathologie, Centre Hospitalo Universitaire Timone, Assistance Publique des Hôpitaux de Marseille, Marseille, France,* [3] *Service des Explorations Fonctionnelles Respiratoires Centre Hospitalo Universitaire Nord, Pôle Cardio-Vasculaire et thoracique, Assistance Publique des Hôpitaux de Marseille, Marseille, France*

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Masaharu Takemura,
Tokyo University of Science, Japan
Steven Wilhelm,
University of Tennessee, Knoxville, United States
Steven M. Short,
University of Toronto Mississauga, Canada

***Correspondence:**
Fabienne Bregeon
fabienne.bregeon@ap-hm.fr

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

Received: *06 December 2017*
Accepted: *27 February 2018*
Published: *21 March 2018*

Citation:
Aherfi S, Nappez C, Lepidi H, Bedotto M, Barassi L, Jardot P, Colson P, La Scola B, Raoult D and Bregeon F (2018) Experimental Inoculation in Rats and Mice by the Giant Marseillevirus Leads to Long-Term Detection of Virus. Front. Microbiol. 9:463. doi: 10.3389/fmicb.2018.00463

The presence of the giant virus of amoeba Marseillevirus has been identified at many different sites on the human body, including in the bloodstream of asymptomatic subjects, in the lymph nodes of a child with adenitis, in one adult with Hodgkin's disease, and in the pharynx of an adult. A high seroprevalence of the Marseillevirus has been recorded in the general population. Whether Marseillevirus can disseminate and persist within a mammal after entry remains unproven. We aimed to assess the ability of the virus to disseminate and persist into healthy organisms, especially in the lymphoid organs. Parenteral inoculations were performed by intraperitoneal injection (in rats and mice) or intravenous injection (in rats). Airway inoculation was performed by aerosolization (in mice). Dissemination and persistence were assessed by using PCR and amebal co-culture. Serologies were performed by immunofluorescent assay. Pathological examination was conducted after standard and immunohistochemistry staining. After intraperitoneal inoculation in mice and rats, Marseillevirus was detected in the bloodstream during the first 24 h. Persistence was noted until the end of the experiment, i.e., at 14 days in rats. After intravenous inoculation in rats, the virus was first detected in the blood until 48 h and then in deep organs with infectious virus detected until 14 and 21 days in the liver and the spleen, respectively. Its DNA was detected for up to 30 days in the liver and the spleen. After aerosolization in mice, infectious Marseillevirus was present in the lungs and nasal associated lymphoid tissue until 30 days post inoculation but less frequently and at a lower viral load in the lung than in the nasal associated lymphoid tissue. No other site of dissemination was found after aerosol exposure. Despite no evidence of disease being observed, the 30-day long persistence of Marseillevirus in rats and mice, regardless of the route of inoculation, supports the hypothesis of an infective potential of the virus in certain conditions. Its constant and long-term detection in nasal associated lymphoid tissue in mice after an aerosol exposure

suggests the involvement of naso-pharyngeal associated lymphoid tissues in protecting the host against environmental Marseillevirus.

Keywords: marseillevirus, experimental infection, murine model, giant viruses, *Megavirales*, NCLDV, pathogenicity

INTRODUCTION

Giant viruses of amoebas were discovered in 2003, with the isolation of *Acanthamoeba polyphaga* Mimivirus by co-culturing on amoeba. *Marseilleviridae* is a new family of amoebal giant viruses defined in 2012 (Colson et al., 2013b). Its founding member is Marseillevirus (Boyer et al., 2009), and in addition 12 other members have been described to date including Senegalvirus, Cannes 8 virus, Fontaine Saint Charles virus, Melbournevirus, Lausannevirus, Tokyovirus, Tunsivirus, Insectomime virus, Brazilian Marseillevirus, Golden Marseillevirus, and Port-Miou virus (Boyer et al., 2009; La Scola et al., 2010; Thomas et al., 2011; Lagier et al., 2012; Aherfi et al., 2013, 2014; Boughalmi et al., 2013; Doutre et al., 2014, 2015; Dornas et al., 2016; Takemura, 2016). Subsequently, contact between giant viruses and humans were suggested. Concordant data argue for the pathogenicity of these viruses, such as mimiviruses-associated pneumonia (La Scola et al., 2005; Raoult et al., 2006; Bousbia et al., 2013; Saadi et al., 2013a,b) or the recently described association between phycodnaviruses and cognitive impairment (Yolken et al., 2014). The presence of giant viruses of amoeba, including those of marseilleviruses within human biological material, was more recently revealed by high throughput metagenomics, confirming contacts between these viruses and humans (Colson et al., 2013a; Rampelli et al., 2016; Verneau et al., 2016).

Senegalvirus was the first marseillevirus to be isolated from human samples, following its serendipitous detection during a microbial metagenomic study conducted on the stools of a healthy Senegalese man (Lagier et al., 2012). In 2013, a metagenomic study further revealed the presence of a substantial number of reads matching the Marseillevirus genome in the viral fraction of healthy blood donors (Popgeorgiev et al., 2013a). Hypotheses were then generated around blood carriage and the blood-borne transmission of Marseillevirus. Furthermore, two seroprevalence studies unexpectedly suggested frequent contacts between humans and Marseillevirus (Mueller et al., 2013; Popgeorgiev et al., 2013b). The detection of giant viruses of amoebae in humans in association with clinical symptoms may be coincidental, but this is nevertheless an emerging issue. A single clinical observation has reported the detection of a marseillevirus in a pathological lymph node of a 11-month-old boy with lymphadenitis (Popgeorgiev et al., 2013c). We subsequently reported the presence of Marseillevirus in the blood and lymph nodes of a patient with Hodgkin's disease (Aherfi et al., 2016a). We also detected Marseillevirus DNA by PCR in two pharyngeal samples collected from a 20-year-old patient presenting neurological disorders at a one-year interval, strongly suggesting the viral persistence of this agent in the tonsils (Aherfi et al., 2016b). To date, no data argued for Marseillevirus propagation in mammal cells and no causal relationship has been established between the presence of this virus and clinical symptoms or diseases observed in these different cases. The only one known host of Marseillevirus that allows a complete lytic cycle is Acanthamoeba cells. To our knowledge, the study of viruses in non host organisms and their interaction remain an unexplored area of virology. Taken together, these findings suggest however that the particles or DNA markers of Marseillevirus may persist during a long period in humans in some cases. Such a hypothesis requires further experimental data.

With this goal, we set up a murine model using rats and mice and different routes of inoculation to assess the dissemination and the persistence of Marseillevirus in mammalian organisms. The aerosol route, we think plausible route of transmission of this waterborne virus (Boyer et al., 2009), was tested first with a special focus on the localization and persistence of the virus in the nasal associated lymphoid tissue (NALT) as an equivalent to the human tonsils. We also tested the intraperitoneal and the intravenous routes in mice and rats.

MATERIALS AND METHODS

Ethics Statements and General Procedures *in Vivo*

For animal studies, the experimental protocols, registered by the "Ministère de l'Enseignement Supérieur et de la Recherche" under reference number 20150528122362 and 2015060517005844, were approved by the Institutional Animal Care and Use Committee of Aix-Marseille University "C2EA-14," France. We used Balb/c mice between 4 and 8 weeks old (Envigo Laboratories, Gannat, France) weighing between 16 and 25 g, and Swiss rats weighing between 330 and 770 g. Animals were housed in individual plastic cages (five mice or two rats per cage) in a ventilated pressurized cabinet (A-BOX 160, Noroit, Rezé, France) with free access to water and standard diet food until the experiment. All animals were housed in protected environmental area and received standard diet including dehydrated rodent feed pellets and sterile water.

Airway inoculation was performed by aerosol delivery using the whole-body inhalation exposure system A4224 (IES, Glas-Col LLC, Terre Haute, USA). Intraperitoneal (IP) and intravenous (IV) inoculations were performed under volatile anesthesia with 5% isoflurane, by percutaneous puncture of the abdomen or injection into the tail vain, respectively.

Control animals received phosphate buffered saline (PBS) via the aerosolized, IP or IV routes according to the same time of exposure or the same volume as infected animals. After inoculations, the animals were transferred into cages and housed in a safety cabinet with food and water *ad libitum*.

Serial blood samples were taken from the IV injected rats over time by tail vein puncture to describe the kinetics of viremia. At the end of the experiments, the rats were euthanized with a lethal dose of thiopental (Panpharma, France) administered intraperitoneally and the mice were euthanized with exsanguination performed under volatile anesthesia. Additional blood and organ samples were collected post-mortem.

Strains, Culture Conditions, and Preparation of Infective Inoculums for Animal Experiments

Marseillevirus strain T19 was co-cultured on axenic *Acanthamoeba castellanii*, in peptone yeast extract broth with glucose medium (PYG). The culture supernatants were then concentrated and purified as previously described and finally washed in PBS (Dornas et al., 2015). The purified virus was aliquoted and stored at −80°C for further use. Ten days before the animals were inoculated, the viable virus was quantified by end point dilution by co-culturing on *A. castellanii*. At this end, serial dilutions of the virus suspension with a dilution factor of 10 were inoculated to amoebas at a concentration of 5.10^5/mL deposited in a 24 well plate. Amoebas were inoculated with each dilution of virus in quadruplicate. The amoeba were checked for lysis 7 days after inoculation. The concentration of the viable virus was those that allowed the amoeba lysis in two wells of the four that were inoculated with this concentration. The concentration of purified viable virus ranged between 7 and 7.5 log units per μL. Purified viruses were diluted in PBS immediately before inoculating the animals, to reach the appropriate inoculums (see "Animal experiments").

Inoculation of Marseillevirus *in Vivo*

For airway inoculation, 81 mice (34 males, 47 females) were aerosol-inoculated with a suspension of PBS containing nine log units of viruses per mL placed into the glass vial for liquid venturi aerosol generation following the manufacturer's recommendations and custom settings. As assessed on animals euthanized just after aerosol exposition ($n = 4$), the initial viral lung inoculum ranged between 4.9 and 5.7 (mean 5.6) log units of viral copies per million murine cells.

For the parenteral inoculations, eight log units of viable virus diluted in 300 μL of PBS were injected to 21 rats IV and 12 rats IP. For mice, seven log units of viable virus diluted in 200 μL of PBS were injected IP ($n = 15$).

Follow-Up and Samplings

After the inoculations, the animals were observed daily for signs of discomfort or illness.

The IP route was assessed for 24 h in mice and until 2 weeks post-inoculation (PI) in a series of rats euthanized at 12 h, and at days 1, 3, 7, 14, and 43 PI. The IV route in rats was assessed until day 43 PI, with evaluation taking place at 12 h, and on days 1, 2 (blood only), 3, 7, 14, 21, 30, and 43 PI. The aerosol route in mice was assessed until 1 month PI with evaluation taking place at 2 h, and on days 1, 7, 14, 21, and 30 PI.

For all animals, the spleen, liver, and blood were collected. In addition, the omentum and mesenteric lymph nodes from IP inoculated rats, the cervical lymph nodes from IV inoculated rats, the lungs, the NALT, and the cervical and tracheal lymph nodes from aerosol inoculated mice were sampled immediately *post mortem*. Spleen weight was immediately recorded and blood was aliquoted for PCR and serology.

To avoid detecting the possible contamination of the external organ with the virus due to the IP inoculation process, the removed abdominal organs were decontaminated in two baths of 70° ethanol and then washed in PBS before culture and PCR processing.

Each freshly sampled organ was separately crushed in PBS for amoebal co-culture and DNA extraction was performed for PCR.

Representative samples of lungs, lymph nodes, NALT, spleen and liver from each evaluation time were fixed in 4% formalin for histological analyses, including a total of 10 spleen and liver samples, 11 lung samples, and 12 NALT samples from aerosol inoculated mice.

Amoebal Co-culture

A. castellanii was cultured at 28°C in PYG. When amoebas were confluent, they were centrifuged, and the pellet was resuspended in sterile Page's amoeba saline solution twice. Finally, amoebas were resuspended in survival buffer solution at a final concentration of between 5.10^5 and 1.10^6 cells/mL with an antimicrobial mix consisting of imipenem / cilastatin (10 μg/mL), vancomycin (10 μg/mL), ciprofloxacin (20 μg/mL), doxycycline (20 μg/mL), and voriconazole (20 μg/mL). Amoebas were distributed in 24-well plates (500 μL of amoeba culture per well). 50 μL of each of the crushed organs was then deposited on the cell layer and incubated at 30°C for 3 days. Two sub-cultures were performed. When amoeba lysis occurred, 100 μL of the well content was spotted on a slide and colored using hemacolor staining (Hemacolor®, Merck, Darmstadt, Germany) to check for the presence of viral factories (Boughalmi et al., 2012; **Figure 1**). Wells containing only amoebas were included in each microplate as negative controls.

Molecular Detection of Marseillevirus

The DNA from the total blood and from the crushed organs was extracted using a QIAamp Tissue Kit (Qiagen). Two systems of specific primers and probes were used for quantitative real-time PCR (reg4-2-F: CCCAACAGAGGCCGAAATT, reg 4-2R: CCTTCTGTACGAGGCCAAAA, probe reg4-2: TCCTCCCCAGAACCAGACTCTCCA, reg 8-2 F: TCT TGTCTGGCTTTCCCTTC, reg8-2 R: GTGTCTCTG CCTGTCCAAA, probe reg8-2: AGTGAGGAGTCTG TTGGCCGCA). These two systems target specifically MAR_ORF210 (encoding a hypothetical protein excluded from Genbank database due to the lack of start codon) and MAR_ORF055 (encoding a RNA polymerase Rpb1 domains 1-2), respectively. These two genes are in single copies in the Marseillevirus strain T19 genome. When an amplification was obtained and a fluorescence signal was generated by testing both the two systems of PCR, the result was considered as positive if the cycle threshold was <35 for at least one of the two systems.

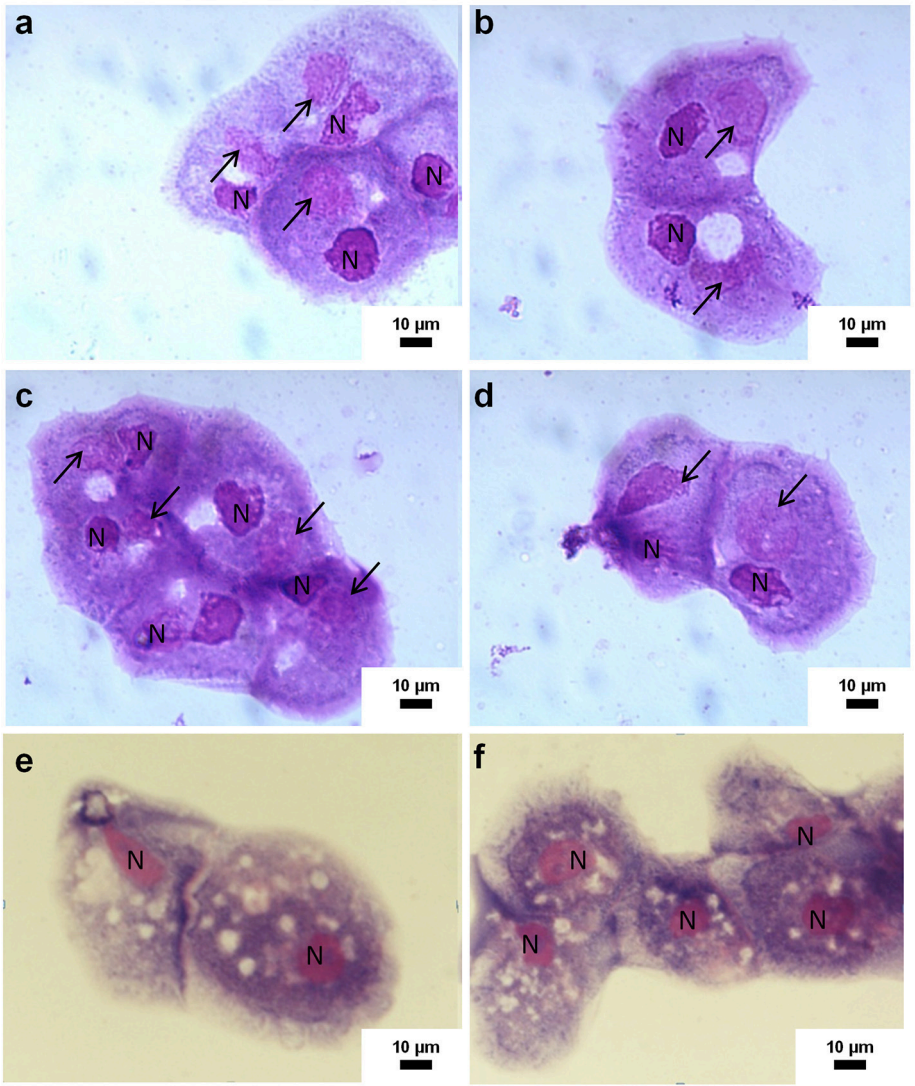


FIGURE 1 | Representative photomicrographs of co-culture of mice and rats samples on *Acanthamoeba castellanii.* Pictures **(a–d)** show positive samples by co-culture as indicated by the visible infected amoebas with Marseillevirus. These samples include NALT from aerosol inoculated mice and spleen from rats after IV inoculation with Marseillevirus. The arrows indicate viral factories and N indicate the nucleus of amoeba. Pictures **(e,f)** show negative samples by amoebal co-culture: the amoebas are not infected and do not contain viral factories.

When amplification was obtained and a fluorescent signal was generated with only one of the two systems, whichever the cycle threshold, the result was considered as negative. The amplification of housekeeping genes hydroxymethylbilane synthase and glyceraldehyde-3-phosphate dehydrogenase were used as internal controls for mice and rats, respectively (Huang et al., 2008; Ding et al., 2014).

Real time PCR assays were performed using the CFX96® qPCR Detection System (Bio-Rad, France). Negative controls consisted of DNA extracted from the organs and blood of PBS-challenged mice and rats (two animals for each route of inoculation). Positive controls were DNA extracted from Marseillevirus culture supernatants.

Viral loads were calculated on the basis of the calibration standard curve of DNA from a suspension of purified Marseillevirus, the concentration of which was determined by flow cytometry (Brussaard, 2004). To standardize the amounts, the viral loads into the tissues were expressed as n log units of viral copies per million murine cells.

Immunofluorescence Assay for Marseillevirus Antibodies Detection in Sera

In the aim to have positive controls for serological tests on the rats and mice of the experiments, we previously immunized a rabbit with Marseillevirus by the subcutaneous route. After three inoculations, serum from the rabbit consisting in polyclonal

antibodies specific to Marseillevirus, was collected and used as a positive control for serological testing.

Purified Marseillevirus was spotted on microscope slides. Sera collected from rats and mice were tested at the 1:50 dilution in PBS. Sera were deposited on the spots and incubated 30 min at 37°C. Slides were washed twice in PBS/Tween20 0.5% during 8 min, once in distilled water during 8 min, then dried. The presence of antibodies was detected using a FITC (fluorescein isothiocyanate) conjugated goat anti-mouse IgG (Immunotech, Marseille, France), anti-mouse IgM at 1: 400 dilution (Jackson Immunoresearch Laboratories, West Grove, USA) for mouse sera, and anti-rat IgG (Jackson ImmunoResearch, Suffolk, United Kingdom) for rat sera, with Evans blue counterstain 0.25%. Slides were incubated at 37°C during 30 min, washed twice in PBS/Tween 20 0.5% during 8 min, once in distilled water during 8 min, then dried. The slides were then observed after adding 1 drop of Fluoprep (Biomérieux, France) and coverslips, on a microscope Leica DM 2,500 (Leica, Wetzlar, Germany) at 488 nm wavelength. As negative controls, sera from non-immunized mice were included in each experiment. Positive controls consisted in sera from immunized rabbit. The threshold for positivity of serology was the 1:50 dilution of the mice and rat sera. A result was considered as positive if the two observers, blind to group assignment, so concluded. Any discordant result was considered as negative.

RESULTS

Dissemination of the Virus

In IP inoculated rats and mice, dissemination of the virus into the bloodstream was observed at day 1 PI, as attested by positive PCR in two of the four rats and two of the four mice tested (Supplementary Files 1, 2). The mean viral loads were at day 0, 3.9, and 6.1 and at day 1, 3.9, and 8 log units per million murine cells, respectively in rats and mice. Blood samples induced amoeba lysis in 5/7 of the rat and mice blood samples collected at day 0 and 2/8 at day 1 PI. Dissemination of the viable virus to deep organs was also observed in the spleen and liver (see below "Persistence of viruses").

As expected, after IV inoculation in rats, the virus was detected in blood samples, but also in deep organs as attested by amoeba co-culture and PCR (**Tables 1**, **2**, **Figure 2**, Supplementary File 3). At 24 h PI, the liver, spleen and lungs were found positive for all the rats tested (4/4).

In aerosol inoculated mice, all animals had negative PCR and co-culture for the blood, spleen, liver and lymph nodes. In contrast, the lungs and NALT were frequently positive, regardless of the sample time, i.e., in 111 of 133 (83%) of the whole tested samples from aerosol inoculated mice, including 50 of 70 (71%) lung samples and 61 of 63 (97%) NALT samples (**Figure 3**, Supplementary File 4).

Persistence of Viruses

After IP inoculation, viable Marseillevirus i.e., detected by co-culture was detected in the spleen from the 12 h PI and persisted until the end of the experiment 14 days later (3/3 rats) (Supplementary File 2). In the liver and the omentum, the viable virus was recovered in the first seven days in rats, and viral DNA i.e., detected by PCR persisting up to day 14 in both organs.

After IV inoculation in rats, the blood detection of Marseillevirus persisted up to 48 h PI in eight of the eight tested blood samples (PCR and culture). In the other organs, the viable virus was detected until days 14 and 21 in the liver and the spleen, respectively, while viral DNA persisted up to day 30 PI in both organs (**Figure 4**, Supplementary File 3).

After aerosolization, viable Marseillevirus persisted at least 30 days in the NALT in all the mice tested, and in only one lung sample of the four collected at the same time point (**Figure 3**, Supplementary File 4). Immediately after aerosolization and after 12 h post exposure, the viral load did not significantly differ between the NALT and the lung samples ($p = 0.27$). Between days 1 and 7 PI, the NALT viral loads increased. Moreover, from days 1 to 21 PI, the NALT viral loads were higher than in the lungs, then, despite decreasing, remained above the lung viral load (**Figure 5**).

Serology

A total of 111 sera including nine from IP inoculated rats, 20 from IV inoculated rats and 82 from aerosol inoculated mice were tested.

After parenteral inoculation, of the eight sera collected between days 1 to 7 PI, two collected on day 7 PI, were positive for anti-Marseillevirus IgG. In IV inoculated rats, IgG anti-Marseillevirus antibodies were found in one of four sera collected on day 7 PI and 10 of 11 sera tested between days 14 and 43 PI. A representative microphotograph is presented in Supplementary File 5.

Only one aerosol inoculated mouse showed an IgG antibody response to Marseillevirus (sampled at day 30 PI).

All sera were negative for IgM, regardless of the inoculation route.

In eight cases, a positive signal was found by only one of the two observers. These samples were recorded as being negative. This concerned IgM antibodies on day 7 for two animals and day 16 for two others, and IgG at day 23 for four animals.

Clinical Outcome

No spontaneous deaths occurred and no animal presented signs of discomfort throughout the course of the experiment, whatever the route of virus inoculation. A regular gain in body weight occurred in all infected and control animals.

Histopathological Findings

No histological lesions were found in any murine tissue including NALT, the lungs, spleen, liver, cervical and tracheal lymph nodes.

DISCUSSION

In this paper, we describe, for the first time to our knowledge, the purposeful transmission of the giant Marseillevirus to a murine host. By including different routes of inoculation, our model aimed to assess the tropism, persistence and dissemination of the virus. We report a 30-day long persistence of the virus in immunocompetent rats and mice inoculated by the IP, IV and respiratory routes. The virus was able to disseminate from the

TABLE 1 | Summary of results obtained by qPCR of blood and organs of rats and mice inoculated with Marseillevirus.

PCR			Day 0	Day 1	Day 2	Day 3	Day 7	Day 14	Day 21	Day 30	Day 43
IP route	Rats	Blood	Positive 33.0 ($n = 2/3$)	Positive 33.3 ($n = 1/4$)	ND	Negative - ($n = 0/2$)	Negative - ($n = 0/4$)	Negative - ($n = 0/3$)	ND	ND	NI
		Spleen	ND	Positive 33.7 ($n = 2/2$)	ND	Positive 34.9 ($n = 1/2$)	Positive 34.7 ($n = 3/3$)	Positive 33.6 ($n = 3/3$)	ND	ND	Negative - ($n = 0/1$)
		Liver	ND	Positive 33.3 ($n = 1/2$)	ND	Positive 34.9 ($n = 1/2$)	Negative ($n = 0/2$)	Positive 33.7 ($n = 2/3$)	ND	ND	Negative - ($n = 0/1$)
	Mice	Blood	Positive 31.8 ($n = 2/4$)	Positive 31.7 ($n = 2/4$)	Positive 32.0 (1/3)	ND	Negative - ($n = 0/3$)	ND	Negative - ($n = 0/1$)	ND	ND
		Spleen	Positive 31.7 ($n = 3/4$)	Positive 33.2 ($n = 2/4$)	Positive 34.2 ($n = 1/3$)	ND	Positive 31.0 ($n = 1/3$)	ND	Negative - ($n = 0/1$)	ND	ND
		Liver	ND	ND	ND	ND	ND	ND	ND	ND	ND
IV route	Rats	Blood	ND	Positive 33.2 ($n = 4/4$)	Positive 32.8 ($n = 2/2$)	Negative - ($n = 0/2$)	Negative - ($n = 0/3$)	Negative - ($n = 0/3$)	Negative - ($n = 0/3$)	Negative - ($n = 0/3$)	Negative - ($n = 0/2$)
		Spleen	ND	Positive 29.4 ($n = 4/4$)	ND	Positive 31.1 ($n = 3/3$)	Positive 29.0 ($n = 2/3$)	Positive 31.5 ($n = 3/3$)	Positive 28.2 ($n = 2/2$)	Positive 28.6 ($n = 2/3$)	Negative - ($n = 0/2$)
		Liver	ND	Positive 29.2 ($n = 4/4$)	ND	Positive 29.0 ($n = 3/3$)	Positive 31.0 ($n = 1/3$)	Positive 30.2 ($n = 3/3$)	Positive 28.2 ($n = 2/2$)	Positive 29.1 ($n = 2/3$)	Negative - ($n = 0/2$)
Aerosol route	Mice	Blood	Negative - ($n = 0/4$)	Negative - ($n = 0/15$)	ND	ND	Negative - ($n = 0/17$)	Negative - ($n = 0/15$)	Negative - ($n = 0/12$)	Negative - ($n = 0/5$)	ND
		Spleen	Negative - ($n = 0/4$)	Negative - ($n = 0/15$)	ND	ND	Negative - ($n = 0/18$)	Negative - ($n = 0/15$)	Negative - ($n = 0/14$)	Negative - ($n = 0/5$)	ND
		Liver	Negative - ($n = 0/4$)	Negative - ($n = 0/15$)	ND	ND	Negative - ($n = 0/18$)	Negative - ($n = 0/15$)	Negative - ($n = 0/14$)	Negative - ($n = 0/5$)	ND
		Lung	Positive 26.1 ($n = 4/4$)	Positive 28.4 ($n = 15/15$)	ND	ND	Positive 30.7 ($n = 16/18$)	Positive 34.6 ($n = 8/15$)	Negative - ($n = 0/14$)	Negative - ($n = 0/4$)	ND
		NALT	Positive 29.7 ($n = 2/2$)	Positive 31.0 ($n = 11/15$)	ND	ND	Positive 31.5 ($n = 18/18$)	Positive 33.6 ($n = 13/15$)	Positive 34.8 ($n = 7/14$)	Positive 33.2 ($n = 3/3$)	ND

Mean Ct obtained by qPCR for IP route and aerosol routes in mice, IP and IV routes in rats. In each case, the first line is the result: positive / negative, the second line is the mean Ct obtained for the positive samples tested, the third line is the number of positive samples/number of tested samples. ND, Not Done.

peritoneum to the bloodstream as well as from the bloodstream into several deep organs. The NALT, a rodent equivalent of the human tonsils, appeared to be an important target organ after aerosol transmission, as attested by its early and lasting carriage at high viral loads as compared to other organs. The viral load, as assessed by quantification of DNA copies when possible, did not increase over the time regardless the animal model or the route of inoculation, so we cannot clearly conclude to the evidence of *in vivo* replication of the virus.

The presence of giant viruses in mammalian hosts was first suggested for mimiviruses, other giant viruses which are close relatives of marseilleviruses. Thus, Mimivirus-associated pneumonia have been described, notably in one patient from which the virus could be isolated from its broncho-alveolar fluid (Saadi et al., 2013a). Another case featuring a laboratory technician handling Mimivirus who developed unexplained pneumonia and seroconversion to Mimivirus antigens which has also been reported (Raoult et al., 2006; Saadi et al., 2013a,b). Moreover, the sero-epidemiological data show a significantly higher seroprevalence for mimivirus in pneumonia patients than in controls. Indeed, on 887 serum samples including 376 from patients with community-acquired pneumonia, and 511 from healthy control subjects, 9.66% of the first group exhibited a positive titer of antibodies to Mimivirus whereas only 2.3% of the

TABLE 2 | Summary of results obtained by amoebal co culture of blood and organs of rats and mice inoculated with Marseillevirus.

Coculture			Day 0	Day 1	Day 2	Day 3	Day 7	Day 14	Day 21	Day 30	Day 43
IP route	Rats	Blood	Positive ($n = 1/2$)	Positive ($n = 1/4$)	ND	Negative ($n = 0/2$)	Negative ($n = 0/4$)	Negative ($n = 0/3$)	ND	ND	Negative ($n = 0/1$)
		Spleen	ND	Positive ($n = 2/2$)	ND	Positive ($n = 1/2$)	Positive ($n = 3/4$)	Positive ($n = 3/3$)	ND	ND	Negative ($n = 0/1$)
		Liver	ND	Positive ($n = 2/2$)	ND	Negative ($n = 0/2$)	Positive ($n = 1/4$)	Negative ($n = 0/3$)	ND	ND	Negative ($n = 0/1$)
	Mice	Blood	Positive ($n = 3/4$)	Positive ($n = 1/4$)	Negative ($n = 0/3$)	ND	Negative ($n = 0/3$)	ND	Negative ($n = 0/1$)	ND	ND
		Spleen	Positive ($n = 4/4$)	Positive ($n = 4/4$)	Positive ($n = 2/3$)	ND	Positive ($n = 2/3$)	ND	Negative ($n = 0/1$)	ND	ND
		Liver	ND	ND	ND	ND	ND	ND	ND	ND	ND
IV route	Rats	Blood	ND	Positive ($n = 4/4$)	Positive ($n = 2/2$)	Negative ($n = 0/3$)	Negative ($n = 0/3$)	Negative ($n = 0/3$)	Negative ($n = 0/3$)	Negative ($n = 0/3$)	Negative ($n = 0/2$)
		Spleen	ND	Positive ($n = 4/4$)	ND	Positive ($n = 2/3$)	Positive ($n = 2/3$)	Positive ($n = 3/3$)	Positive ($n = 2/2$)	Negative ($n = 0/3$)	Negative ($n = 0/2$)
		Liver	ND	Positive ($n = 4/4$)	ND	Positive ($n = 3/3$)	Positive ($n = 2/3$)	Positive ($n = 2/3$)	Negative ($n = 0/2$)	Negative ($n = 0/3$)	Negative ($n = 0/2$)
Aerosol route	Mice	Blood	Negative ($n = 0/4$)	Negative ($n = 0/5$)	ND	ND	Negative ($n = 0/18$)	Negative ($n = 0/15$)	Negative ($n = 0/14$)	Negative ($n = 0/5$)	ND
		Spleen	Negative ($n = 0/4$)	Negative ($n = 0/5$)	ND	ND	Negative ($n = 0/18$)	Negative ($n = 0/15$)	Negative ($n = 0/14$)	Negative ($n = 0/5$)	ND
		Liver	Negative ($n = 0/4$)	Negative ($n = 0/5$)	ND	ND	Negative ($n = 0/18$)	Negative ($n = 0/15$)	Negative ($n = 0/14$)	Negative ($n = 0/5$)	ND
		Lung	Positive ($n = 4/4$)	Positive ($n = 15/15$)	ND	ND	Positive ($n = 16/18$)	Positive ($n = 8/15$)	Positive ($n = 1/14$)	Positive ($n = 1/4$)	ND
		NALT	Positive ($n = 2/2$)	Positive ($n = 11/11$)	ND	ND	Positive ($n = 18/18$)	Positive ($n = 14/15$)	Positive ($n = 11/14$)	Positive ($n = 3/3$)	ND

For some times of evaluation, the number of samples tested by coculture is different from those appearing in the table showing PCR results because of either insufficient quantity or ininterpretable results of PCR (negativity for internal control DNA and Marseillevirus DNA). In each case, the second line is the number of positive samples / number of tested samples. ND = Not done.

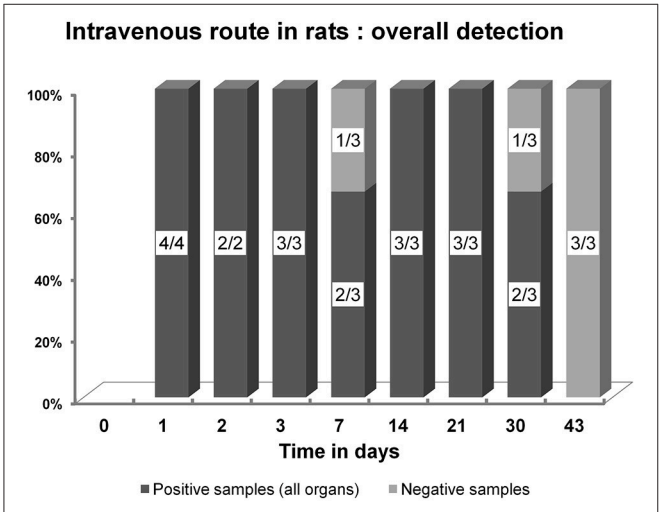


FIGURE 2 | Overall detection and persistence of Marseillevirus DNA by PCR after an intraveinous inoculation in rats, during the 43 day long experiment.

healthy controls were positive ($p = 0.01$; La Scola et al., 2005). Moreover, Mimivirus DNA was detected by PCR in respiratory samples from a patient with hospital-acquired pneumonia (La Scola et al., 2005). However, studies using PCR assays were more difficult to conduct because of the great genetic variability of the mimiviruses genomes, a feature shared with marseilleviruses. Thus, Dare et al. screened 496 respiratory specimens from nine pneumonia patient populations for Mimivirus by qPCR, performed mainly on nasal and nasopharyngeal swabs. All the samples tested were negative (Dare et al., 2008).

The clinical data mentioned above were completed by a mouse model reproducing histologically proven pneumonia at days 3 and 7 PI in C57BL/6 and BALB/c mice respectively (Khan et al., 2007). Another giant virus, *Acanthocystis turfacea* Chlorella Virus 1, a close relative of amoeba giant viruses from the family *Phycodnaviridae*, was found in oro-pharyngeal samples from patients and was associated with a decrease in cognitive functioning (Yolken et al., 2014). A mouse model showed that digestive inoculation of the virus induced, modifications in the brain of the expression of genes involved in cognitive functions. These authors supposed that the virus was responsible for cognitive impairment, although such a hypothesis would need further investigation (Yolken et al., 2014).

In the present work, the IP model showed an early transient blood dissemination of the virus both in mice and rats, and its persistence in the spleen for at least 2 weeks. The IV model also showed that after a transient passage in the bloodstream,

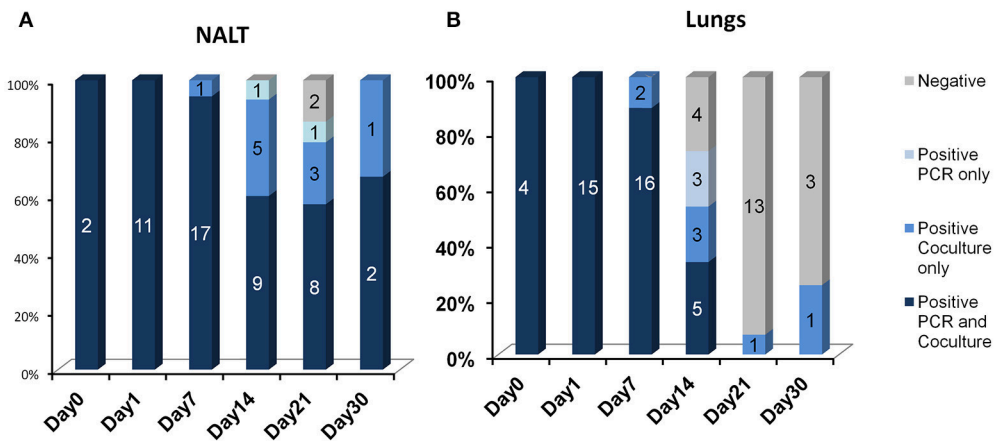


FIGURE 3 | Positive **(A)** NALT and **(B)** lung samples, according to the technique (PCR or amoebal co-culture), after pulmonary inoculation. The percentage of positive samples is indicated on the y axis. The absolute numbers of positive samples are indicated by labels on the plots.

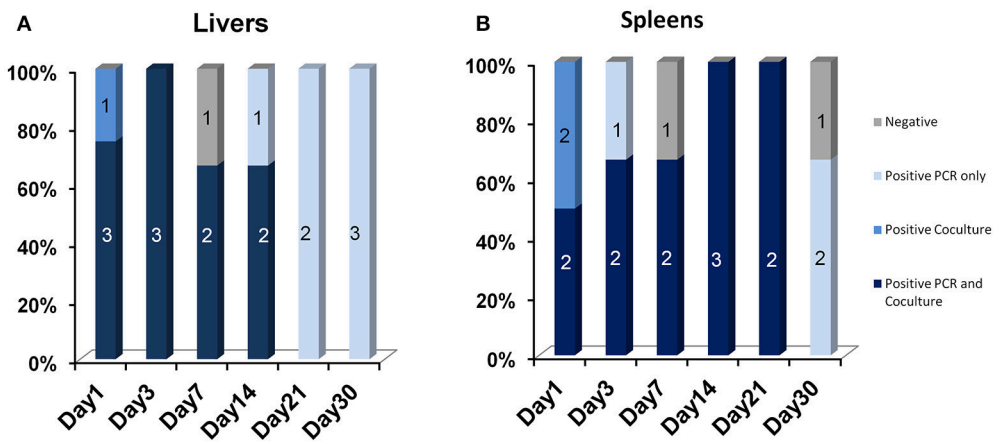


FIGURE 4 | Positive **(A)** liver and **(B)** spleen samples, according to the technique (PCR or amoebal co-culture), after intravenous inoculation in rats. The percentage of positive samples is indicated on the y axis. The absolute numbers of positive samples are indicated by labels on the plots.

viable Marseillevirus was detected as much as 3 weeks later in the spleen. In the aerosolized model, the virus was detected at a higher frequency in NALT than in lung samples, especially at later time points. Interestingly, the DNA viral load at day 30 PI was 4.7 log units of viral copies per million of murine cells, in other words, not that different to the load just after aerosolization (5.9 log units of viral copies per million of murine cells). Conversely, in the lungs, the viral load regularly decreased until it was undetectable at day 30 PI. Although our results do not show a viral replication, the long persistence into the NALT of aerosol inoculated mice is congruent with the human case of Marseillevirus persistence in pharyngeal samples (Aherfi et al., 2016b).

The use of two techniques (amoebal co-culture and PCR) for detecting the virus, complemented with antibody detection assays under strict control conditions and predefined strict criteria for the PCR and serology interpretations strengthens our results. In addition, double blind reading of immunofluorescence assays was performed. This could have led to the under diagnosis of positive serological responses after aerosolization. Concerning the antibody response after parenteral inoculations, a strong concordance was obtained.

The absence of any pathological findings in the organs, including the lungs, could be due to an absence of a detectable host cellular immune response or to the invasiveness of the

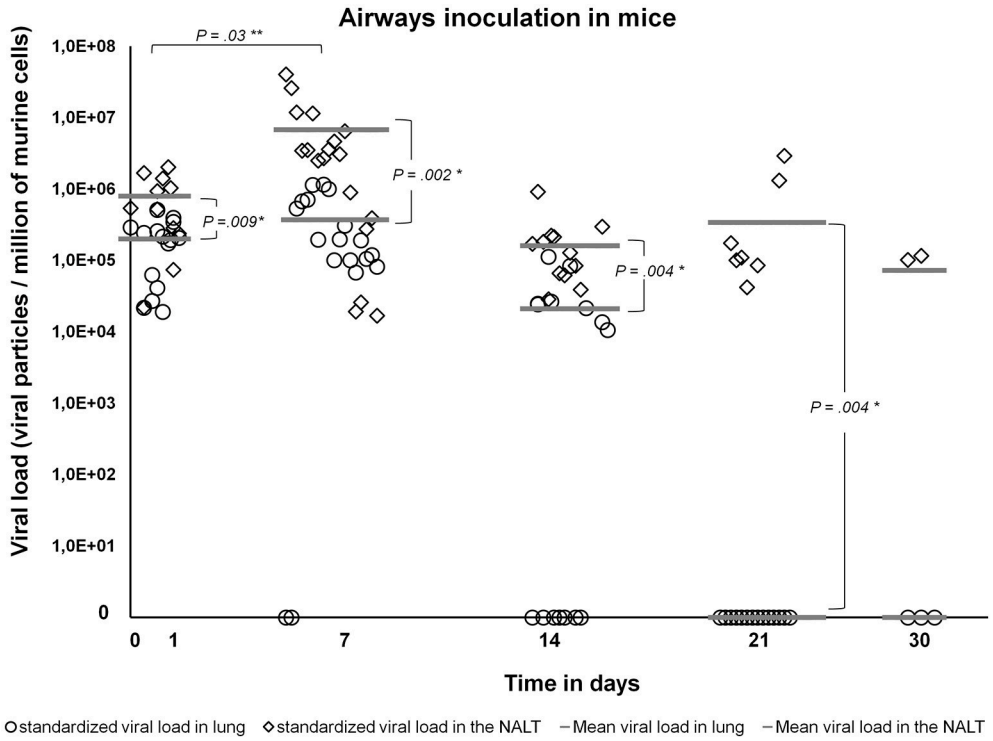


FIGURE 5 | Standardized viral loads (viral copies per million murine cells) in the lungs and the NALT of aerosol-inoculated mice by Marseillevirus suspension. Diamonds represent the viral loads in the NALT and circles represent the viral loads in the lungs for each animal. The mean viral loads are represented for each time point by thick dashes. The viral loads in lungs and NALT, and at different time points were compared by using Wilcoxon-Mann Whitney test, using sigmaplot 13 SYSTAT software Inc. software. Single asterisks show the time points with significant differences between NALT and lung viral loads. Double asterisks show significant differences between the NALT viral loads at different time points.

pathogen. However, it is not known whether immune-suppressed animals or repetitive contact may have induced some of these cases.

The presence of marseilleviruses in humans has previously been reported from different cases, including blood from healthy donors, one case of adenitis, one case of Hodgkin's lymphoma and, as a chronic carriage, in a patient with neurological symptoms (Popgeorgiev et al., 2013a,c; Aherfi et al., 2016a,b). Our results in mice and rats reinforce the hypothesis of chronic carriage. Considering the low number of proportion of positive clinical samples, either for mimiviruses or for marseilleviruses, we can hypothesize that the techniques used to detect giant viruses lack of sensitivity. There are undeniably, a lot of technique improvements that remain to do, both on culture isolation and PCR techniques for detecting giant viruses in clinical samples. Thus, the low number of viral particles combined with the lack of sensitivity of the techniques used may lead to a low number of positive samples in the samples collected. It is noteworthy that it has not been established to date that marseilleviruses can replicate in mammals, or cause a disease. However, giant viruses of amoebae, which are very distant from other viruses both by their phenotypic and genotypic features, might act on mammal cells by a different mechanism than replication. Thus, the big size of giant viruses may probably enable their ingestion by phagocytic cells, without the intervention of a specific cell receptor (Ghigo et al., 2008).

The absence of pathological findings in the tested organs points toward the healthy carriage of the virus by the host. However, further investigations should be performed to assess whether recurrent contact with Marseillevirus or if an inoculation in immunocompromised mice may favor a pathologic outcome. Although the virus was not detected in the lymph nodes in our work, it was found to be viable for as long as 2–3 weeks in the spleen after IP and IV inoculation, respectively.

Given the high prevalence of marseilleviruses in the environment, and the possibility of a long term carriage, further investigations are needed on the mechanisms used by these viruses to escape rapid destruction by immune system. It would be interesting to try culturing Marseillevirus on different professional phagocytic cells, as was performed for Mimivirus, to assess if at least some of them are permissive. To date, only amoebas are known to be a host for Marseillevirus that allow a complete lytic replication cycle. However, if humans are possible carriers of Marseillevirus, they might serve as vectors for their dissemination in the environment. In summary, this experimental model is a first step toward the assessment of Marseillevirus infection in a mammalian host. Its long persistence, especially in the NALT, merits further study to assess the possibility of a longer viral persistence and reinforces

the pertinence of systematic Marseillevirus detection in subjects presenting with unexplained upper airway/pharyngeal or adenitis clinical pictures.

AUTHOR CONTRIBUTIONS

DR, FB, and SA designed the project. CN, FB, and SA implemented the animal experiments. HL analyzed the tissue sections for anatomo-pathology. MB set the protocol of PC, LB, and SA performed the amoebal co-cultures. PJ and SA performed the PC experiments. PC, BS, and DR supervised the project. FB and SA wrote the manuscript.

FUNDING

This work was supported by a help from the French State managed by the National Research Agency under the Investissements d'avenir (Investments for the Future) program with the reference ANR-10-IAHU-03 (Méditerranée Infection).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.00463/full#supplementary-material

REFERENCES

Aherfi, S., Boughalmi, M., Pagnier, I., Fournous, G., La Scola, B., Raoult, D., et al. (2014). Complete genome sequence of Tunisvirus, a new member of the proposed family Marseilleviridae. *Arch. Virol.* 159, 2349–2358. doi: 10.1007/s00705-014-2023-5

Aherfi, S., Colson, P., Audoly, G., Nappez, C., Xerri, L., Valensi, A., et al. (2016a). Marseillevirus in lymphoma: a giant in the lymph node. *Lancet Infect. Dis.* 16, e225–e234. doi: 10.1016/S1473-3099(16)30051-2

Aherfi, S., Colson, P., and Raoult, D. (2016b). Marseillevirus in the pharynx of a patient with neurologic disorders. *Emerg. Infect. Dis.* 22, 2008–2010. doi: 10.3201/eid2211.160189

Aherfi, S., Pagnier, I., Fournous, G., Raoult, D., La Scola, B., and Colson, P. (2013). Complete genome sequence of Cannes 8 virus, a new member of the proposed family "Marseilleviridae". *Virus Genes* 47, 550–555. doi: 10.1007/s11262-013-0965-4

Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D., and La Scola, B. (2013). First isolation of a marseillevirus in the diptera *Syrphidae Eristalis* tenax. *Intervirology* 56, 386–394. doi: 10.1159/000354560

Boughalmi, M., Saadi, H., Pagnier, I., Colson, P., Fournous, G., Raoult, D., et al. (2012). High-throughput isolation of giant viruses of the mimiviridae and marseilleviridae families in the tunisian environment. *Environ. Microbiol.* 15, 2000–2007. doi: 10.1111/1462-2920.12068

Bousbia, S., Papazian, L., Saux, P., Forel, J. M., Auffray, J. P., Martin, C., et al. (2013). Serologic prevalence of amoeba-associated microorganisms in intensive care unit pneumonia patients. *PLoS ONE* 8:e58111. doi: 10.1371/journal.pone.0058111

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21848–21853. doi: 10.1073/pnas.0911354106

Brussaard, C. P. (2004). Optimization of procedures for counting viruses by flow cytometry. *Appl. Environ. Microbiol.* 70, 1506–1513. doi: 10.1128/AEM.70.3.1506-1513.2004

Colson, P., Fancello, L., Gimenez, G., Armougom, F., Desnues, C., Fournous, G., et al. (2013a). Evidence of the megavirome in humans. *J. Clin. Virol.* 57, 191–200. doi: 10.1016/j.jcv.2013.03.018

Colson, P., Pagnier, I., Yoosuf, N., Fournous, G., La Scola, B., and Raoult, D. (2013b). "Marseilleviridae," a new family of giant viruses infecting amoebae. *Arch. Virol.* 158, 915–920. doi: 10.1007/s00705-012-1537-y

Dare, R. K., Chittaganpitch, M., and Erdman, D. D. (2008). Screening pneumonia patients for mimivirus. *Emerg. Infect. Dis.* 14, 465–467. doi: 10.3201/eid1403.071027

Ding, S., Fan, Y., Zhao, N., Yang, H., Ye, X., He, D., et al. (2014). High-fat diet aggravates glucose homeostasis disorder caused by chronic exposure to bisphenol A. *J. Endocrinol.* 221, 167–179. doi: 10.1530/JOE-13-0386

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

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

Doutre, G., Arfib, B., Rochette, P., Claverie, J. M., Bonin, P., and Abergel, C. (2015). Complete genome sequence of a new member of the Marseilleviridae recovered from the brackish submarine spring in the cassis port-miou Calanque, France. *Genome Announc.* 3:01148-15. doi: 10.1128/genomeA.01148-15

Doutre, G., Philippe, N., Abergel, C., and Claverie, J. M. (2014). Genome analysis of the first Marseilleviridae representative from Australia indicates that most of its genes contribute to virus fitness. *J. Virol.* 88, 14340–14349. doi: 10.1128/JVI.02414-14

Ghigo, E., Kartenbeck, J., Lien, P., Pelkmans, L., Capo, C., Mege, J. L., et al. (2008). Ameobal pathogen mimivirus infects macrophages through phagocytosis. *PLoS Pathog.* 4:e1000087. doi: 10.1371/journal.ppat.1000087

Huang, X. Y., Chen, F. H., Li, J., Xia, L. J., Liu, Y. J., Zhang, X. M., et al. (2008). Mechanism of fibroblast-like synoviocyte apoptosis induced by recombinant human endostatin in rats with adjuvant arthritis. *Anat. Rec.* 291, 1029–1037. doi: 10.1002/ar.20722

Khan, M., La Scola, B., Lepidi, H., and Raoult, D. (2007). Pneumonia in mice inoculated experimentally with Acanthamoeba polyphaga mimivirus. *Microb. Pathog.* 42, 56–61. doi: 10.1016/j.micpath.2006.08.004

Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., et al. (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* 18, 1185–1893. doi: 10.1111/1469-0691.12023

La Scola, B., Campocasso, A., N'Dong, R., Fournous, G., Barrassi, L., Flaudrops, C., et al. (2010). Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. *Intervirology* 53, 344–353. doi: 10.1159/000312919

La Scola, B., Marrie, T. J., Auffray, J. P., and Raoult, D. (2005). Mimivirus in pneumonia patients. *Emerg. Infect. Dis.* 11, 449–452. doi: 10.3201/eid1103.040538

Mueller, L., Baud, D., Bertelli, C., and Greub, G. (2013). Lausannevirus seroprevalence among asymptomatic young adults. *Intervirology* 56, 430–433. doi: 10.1159/000354565

Popgeorgiev, N., Boyer, M., Fancello, L., Monteil, S., Robert, C., Rivet, R., et al. (2013a). Marseillevirus-like virus recovered from blood donated by asymptomatic humans. *J. Infect. Dis.* 208, 1042–1050. doi: 10.1093/infdis/jit292

Popgeorgiev, N., Colson, P., Thuret, I., Chiarioni, J., Gallian, P., Raoult, D., et al. (2013b). Marseillevirus prevalence in multitransfused patients suggests blood transmission. *J. Clin. Virol.* 58, 722–725. doi: 10.1016/j.jcv.2013.10.001

Popgeorgiev, N., Michel, G., Lepidi, H., Raoult, D., and Desnues, C. (2013c). Marseillevirus adenitis in an 11-month-old child. *J. Clin. Microbiol.* 51, 4102–4105. doi: 10.1128/JCM.01918-13

Rampelli, S., Soverini, M., Turroni, S., Quercia, S., Biagi, E., Brigidi, P., et al. (2016). ViromeScan: a new tool for metagenomic viral community profiling. *BMC Genomics* 17:165. doi: 10.1186/s12864-016-2446-3

Raoult, D., Renesto, P., and Brouqui, P. (2006). Laboratory infection of a technician by mimivirus. *Ann. Intern. Med.* 144, 702–703. doi: 10.7326/0003-4819-144-9-200605020-00025

Saadi, H., Pagnier, I., Colson, P., Cherif, J. K., Beji, M., Boughalmi, M., et al. (2013a). First isolation of Mimivirus in a patient with pneumonia. *Clin. Infect. Dis.* 57, e127–e134. doi: 10.1093/cid/cit354

Saadi, H., Reteno, D. G., Colson, P., Aherfi, S., Minodier, P., Pagnier, I., et al. (2013b). Shan virus: a new mimivirus isolated from the stool of a Tunisian patient with pneumonia. *Intervirology* 56, 424–429. doi: 10.1159/000354564

Takemura, M. (2016). Draft genome sequence of tokyovirus, a member of the family marseilleviridae isolated from the arakawa river of Tokyo, Japan. *Genome Announc.* 4, e00429–e00416. doi: 10.1128/genomeA.00429-16

Thomas, V., Bertelli, C., Collyn, F., Casson, N., Telenti, A., Goesmann, A., et al. (2011). Lausannevirus, a giant amoebal virus encoding histone doublets. *Environ. Microbiol.* 13, 1454–1466. doi: 10.1111/j.1462-2920.2011.02446.x

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

Yolken, R. H., Jones-Brando, L., Dunigan, D. D., Kannan, G., Dickerson, F., Severance, E., et al. (2014). Chlorovirus ATCV-1 is part of the human oropharyngeal virome and is associated with changes in cognitive functions in humans and mice. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16106–16111. doi: 10.1073/pnas.1418895111

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 27 November 2018
doi: 10.3389/fmicb.2018.02668

Ancestrality and Mosaicism of Giant Viruses Supporting the Definition of the Fourth TRUC of Microbes

*Philippe Colson[1], Anthony Levasseur[1], Bernard La Scola[1], Vikas Sharma[1,2], Arshan Nasir[3,4], Pierre Pontarotti[1,2], Gustavo Caetano-Anollés[3] and Didier Raoult[1]**

[1] Aix-Marseille Université, Institut de Recherche pour le Développement (IRD), Assistance Publique – Hôpitaux de Marseille (AP-HM); Microbes, Evolution, Phylogeny and Infection (MEΦI); Institut Hospitalo-Universitaire (IHU) – Méditerranée Infection, Marseille, France, [2] Centre National de la Recherche Scientifique, Marseille, France, [3] Evolutionary Bioinformatics Laboratory, Department of Crop Sciences, University of Illinois Urbana-Champaign, Urbana, IL, United States, [4] Department of Biosciences, COMSATS University Islamabad, Islamabad, Pakistan

OPEN ACCESS

Edited by:
Steven M. Short, University of Toronto Mississauga, Canada

Reviewed by:
Jessica Labonté, Texas A&M University at Galveston, United States
Jozef I. Nissimov, Rutgers, The State University of New Jersey, United States
David Robert Wessner, Davidson College, United States

***Correspondence:**
Didier Raoult
didier.raoult@gmail.com

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

Received: *22 May 2018*
Accepted: *18 October 2018*
Published: *27 November 2018*

Citation:
Colson P, Levasseur A, La Scola B, Sharma V, Nasir A, Pontarotti P, Caetano-Anollés G and Raoult D (2018) Ancestrality and Mosaicism of Giant Viruses Supporting the Definition of the Fourth TRUC of Microbes. Front. Microbiol. 9:2668. doi: 10.3389/fmicb.2018.02668

Giant viruses of amoebae were discovered in 2003. Since then, their diversity has greatly expanded. They were suggested to form a fourth branch of life, collectively named 'TRUC' (for "Things Resisting Uncompleted Classifications") alongside *Bacteria*, *Archaea,* and *Eukarya*. Their origin and ancestrality remain controversial. Here, we specify the evolution and definition of giant viruses. Phylogenetic and phenetic analyses of informational gene repertoires of giant viruses and selected bacteria, archaea and eukaryota were performed, including structural phylogenomics based on protein structural domains grouped into 289 universal fold superfamilies (FSFs). Hierarchical clustering analysis was performed based on a binary presence/absence matrix constructed using 727 informational COGs from cellular organisms. The presence/absence of 'universal' FSF domains was used to generate an unrooted maximum parsimony phylogenomic tree. Comparison of the gene content of a giant virus with those of a bacterium, an archaeon, and a eukaryote with small genomes was also performed. Overall, both cladistic analyses based on gene sequences of very central and ancient proteins and on highly conserved protein fold structures as well as phenetic analyses were congruent regarding the delineation of a fourth branch of microbes comprised by giant viruses. Giant viruses appeared as a basal group in the tree of all proteomes. A pangenome and core genome determined for *Rickettsia bellii* (bacteria), *Methanomassiliicoccus luminyensis* (archaeon), *Encephalitozoon intestinalis* (eukaryote), and Tupanvirus (giant virus) showed a substantial proportion of Tupanvirus genes that overlap with those of the cellular microbes. In addition, a substantial genome mosaicism was observed, with 51, 11, 8, and 0.2% of Tupanvirus genes best matching with viruses, eukaryota, bacteria, and archaea, respectively. Finally, we found that genes themselves may be subject to lateral sequence transfers. In summary, our data highlight the quantum leap between classical and giant viruses. Phylogenetic and phyletic analyses and the study of protein fold superfamilies confirm previous evidence

of the existence of a fourth TRUC of life that includes giant viruses, and highlight its ancestrality and mosaicism. They also point out that best evolutionary representations for giant viruses and cellular microorganisms are rhizomes, and that sequence transfers rather than gene transfers have to be considered.

Keywords: giant virus, TRUC, megavirales, mimivirus, informational genes, protein structural domains

INTRODUCTION

Since the Mimivirus discovery in 2003, dozens of giant viruses that infect *Acanthamoeba* spp. or *Vermamoeba vermiformis* have been isolated from various environmental samples, and more recently from animals including humans (La Scola et al., 2003; Raoult et al., 2004; Colson et al., 2017a). Currently, families *Mimiviridae* (La Scola et al., 2005) and *Marseilleviridae* (Boyer et al., 2009; Colson et al., 2013b) and isolates that represent new putative families of giant viruses of amoebae, including pandoraviruses (Philippe et al., 2013), pithoviruses (Legendre et al., 2015), faustoviruses (Reteno et al., 2015), Mollivirus (Legendre et al., 2015), Kaumoebavirus (Bajrai et al., 2016), cedratviruses (Andreani et al., 2016), Pacmanvirus (Andreani et al., 2017), and Orpheovirus (Andreani et al., 2018) have been described (Colson et al., 2017b). These giant viruses of amoebae exhibit unique phenotypic and genotypic characteristics that differentiate them from 'traditional' viruses and bring them close to small microbes (Lwoff, 1957; Colson et al., 2017a).

These viruses were linked through phylogenomic analyses to poxviruses, asfarviruses, ascoviruses, iridoviruses, and phycodnaviruses (formerly the largest viral representatives), which were grouped in 2001 in a superfamily named nucleocytoplasmic large DNA viruses (NCLDVs) (Iyer et al., 2001, 2006; Raoult et al., 2004). NCLDVs and giant viruses of amoebae were reported to share a putative ancient common ancestor harboring about 50 conserved core genes responsible for key viral functions (Yutin et al., 2009; Koonin and Yutin, 2010; Yutin and Koonin, 2012). Together with a common virion architecture and common major biological features including reproduction within cytoplasmic factories, this contributed to propose reclassifying NCLDVs, mimiviruses and marseilleviruses in a new viral order named Megavirales (Colson et al., 2013a).

The origin and ancestrality of giant viruses has remained controversial. From the onset, when the Mimivirus genome was sequenced in 2004, a phylogeny based on seven concatenated universally conserved genes showed that Mimivirus branched near the origin of the eukaryotic branch, and it was suggested that giant viruses comprised a fourth additional branch in the Tree of Life, alongside *Bacteria*, *Archaea,* and *Eukarya* (Raoult et al., 2004). This hypothesis was thereafter strengthened by both cladistic and phenetic analyses based on informational genes, including those implicated in nucleotide biosynthesis, transcription and translation (Boyer et al., 2010). The hypothesis of the existence of a fourth branch of microbes prompted to define the 'TRUCs,' which is an acronym for "Things Resisting Uncompleted Classifications" (Raoult, 2013, 2014). This term was coined because the definition of domains of life by C.R. Woese was based on ribosomal genes that are absent in giant viruses. This proposal of a fourth branch of life comprised by giant viruses has remained controversial and a subject of debate among virologists and evolutionary biologists. Some phylogenetic analyses were deemed to suggest complex patterns of evolutionary relationships for different informational proteins from giant viruses, which even questioned the monophyly of NCLDVs (Yutin and Koonin, 2012; Yutin et al., 2014). A high level of mosaicism has been highlighted for the genomes of giant viruses of amoebae, which was related to sequence transfers with organisms belonging to the three cellular domains of Life (Raoult et al., 2004; Boyer et al., 2009). A substantial gene flow has been also described in NCLDVs including in coccolithoviruses (Wilson et al., 2009; Nissimov et al., 2017). It was suspected that lateral gene transfers blurred phylogenies based on genes shared by giant viruses and cellular organisms (Moreira and Lopez-Garcia, 2009). Several phylogenetic reconstructions in which giant viruses branch within eukaryotes were published (Moreira and Lopez-Garcia, 2009, 2015; Williams et al., 2011), and it was put forward that the universally conserved genes used in phylogeny reconstructions might have been acquired by giant viruses from their proto-eukaryotic hosts (Moreira and Lopez-Garcia, 2009; Yutin et al., 2014). The interpretation of some phylogenies was also that modern giant viruses might originate from smaller NCLDVs (Yutin and Koonin, 2013; Yutin et al., 2014). Conversely, it was proposed that giant viruses might derive from ancestral cellular genomes by reductive evolution (Legendre et al., 2012). Besides, phylogenetic reconstructions supporting the fourth TRUC hypothesis triggered methodological criticisms arguing that they were distorted by long-branch attraction and technical issues, and divergences in their interpretation. However, alternative phylogenies were not accurate either regarding the phylogeny of *Archaea*, *Bacteria,* or *Eukarya* (Williams et al., 2011; Moreira and Lopez-Garcia, 2015). A four-branch topology was also obtained by reconstructing phylogenies that describe the evolution of proteomes and protein domain structures (Nasir et al., 2012; Nasir and Caetano-Anollés, 2015). The genomic and structural diversity embedded in giant virus proteomes was found similar to that of proteomes of cellular organisms with parasitic lifestyles. Beyond, other phylogenies based on RNA polymerase suggested the presence in metagenomes of sequences related to giant virus relatives (Wu et al., 2011; Sharma et al., 2014). As a synthesis, it was deemed that more work is needed on Megavirales phylogenies to clarify if these viruses are monophyletic or have different evolutionary histories (Forterre and Gaia, 2016). Here, we specify the definition of giant viruses, highlight their mosaicism at the genome, structure and sequence level, and strengthen the

evidence for their ancestrality and the existence of a fourth TRUC of microbes.

MATERIALS AND METHODS

Definition of Giant Viruses

We collected and reviewed current knowledge on giant viruses from articles gathered from the NCBI PubMed database and from Google Scholar using as keywords "giant virus"; Megavirales; mimivir*; marseillevir*; pandoravir*; pithovir*; faustovir*; mollivirus; cedratvirus; kaumoebavirus; pacmanvirus; virophage; transpoviron. We then compared the phenotypic and genotypic features of these viruses with those used as criteria to define classical viruses and those that are hallmark features of cellular organisms. The list of those criteria is presented in **Table 1**.

Protein Structure Assignment to Viral and Cellular Proteomes

Protein sequences from completely-sequenced proteomes of 80 Megavirales were scanned against the library of hidden Markov models (HMMs) of structural recognition maintained by the SUPERFAMILY database for structure assignment at an *E*-value cutoff of < 0.0001 (Gough et al., 2001; Gough and Chothia, 2002). The SUPERFAMILY HMMs represent proteins of known three-dimensional (3D) structures and assign each detected occurrence of protein domain into fold superfamilies (FSFs), as defined by the Structural Classification of Proteins (ver. 1.75) database (Andreeva et al., 2008). FSFs are collections of one or more protein families that show recognizable 3D structural and functional similarities, but not necessarily sequence identities, that are indicative of common origin. Thus, FSFs represent highly dissimilar protein domains at the sequence level that have evolved via divergence from a common structure and can still be recognized based on the presence of that conserved structural core by HMMs trained to detect remote homologies. Because of the fast mutation rates of viral genes, it sometimes becomes impossible to generate meaningful global sequence alignments when considering viral and cellular genes together in data matrices. The fast mutation rates, especially when considering proteins separated by large evolutionary distances and involving distantly related taxa, lead to alignment inaccuracies and large number of gaps. In contrast, protein structure evolves at least 3 to 10 times slower than molecular sequences (Illergard et al., 2009) and hence provides an alternative to study the deep evolutionary history of cells and viruses (Nasir et al., 2012; Nasir and Caetano-Anollés, 2015). In parallel, FSF assignments for a total of 102 cellular organisms including an equal number of archaea, bacteria, and eukaryota were retrieved from a previous work during which a total of 1,797 distinct FSF domains had been detected (*E*-value < 0.0001) (Nasir and Caetano-Anollés, 2015).

Structural Phylogenomics

Using an in-house Python script, we generated a data matrix containing 182 rows (proteomes from 34 archaea, 34 bacteria, 34 eukaryota, and 80 Megavirales members) and 289 columns (FSFs) containing presence/absence information for 'universal' FSFs. 'Universal' FSFs, by definition, included FSFs that were detected in at least one proteome each from archaea, bacteria, eukaryota, and a Megavirales member. In other words, FSFs unique to one of these four groups (e.g., bacteria-specific FSFs) or shared by 2-to-3 groups of cellular organisms and/or viruses (e.g., FSFs detected in archaea, bacteria, and viruses but not eukaryota) were excluded from our definition of universal FSFs (see (Nasir et al., 2015) for details on FSF groups in cellular organisms and viruses). This data matrix containing 182 proteomes and 289 universal FSFs was imported into the PAUP (ver. 4.0b10) software (Swofford, 2002) for phylogenomic tree reconstruction. Proteomes were treated as taxa and FSFs as characters. Presence/absence of FSFs (represented by 1 and 0, respectively) were used as distinct character states to distinguish taxa. Maximum parsimony method was set as optimality criterion to reconstruct the most parsimonious unrooted phylogenomic tree describing the evolution of sampled proteomes based on the presence/absence of 286 parsimony informative FSF characters. The unrooted reconstructed tree was rooted *a posteriori* by the branch resulting in minimum increase in overall tree length using the Lundberg method (Lundberg, 1972; see Nasir et al., 2017; Caetano-Anollés et al., 2018 for description and review of rooting methodology). The reliability of the phylogenetic splits was evaluated by running 1,000 bootstraps. Separately, we performed principal coordinate analysis (PCoA) on the same data matrix and plotted the 182 sampled viral and cellular proteomes into 3D space. Proteomes are composed of FSF domains of different evolutionary and geological ages. From a previously reconstructed tree of domains (ToD) (Nasir and Caetano-Anollés, 2015), we retrieved the relative evolutionary ages for each of the 289 universal FSFs. The relative scale reflects the distance of each node (FSF domain) from the root of the ToD and ranges from 0 (closer to the root, most ancient) to 1 (most recent). The node distance (*nd*) value thus describes a clock-like behavior for the evolution of FSF domains and has previously been linked to the geological record (Wang et al., 2011). Euclidean distance was used to plot proteome dissimilarity based on the 1-*nd* transformation of the *nd* scale for each FSF domain in every proteome, as previously (Nasir and Caetano-Anollés, 2015). Since the PCoA is centered around *nd* variable derived from an evolutionary tree, we refer to this method as evo-PCoA. The evo-PCoA thus projects proteome dissimilarity into 3D space based on differences in the evolutionary ages of components of each proteome. XLSTAT plugin was added to Microsoft Excel for generation of PCoA.

Collection of Orthologous Sequences From Viruses

Analysis was performed as described in previous works (Boyer et al., 2010; Sharma et al., 2014). The genes used in the present study were identified from clusters of orthologous groups of proteins (COGs) involved in nucleotide transport and metabolism and information storage and processing (i.e., categories F, J, A, K, L, and B). These genes comprise proteins that

TABLE 1 | Comparison of major features used as criteria to define classical viruses with those of giant viruses and to hallmark features of cellular microbes.

Phenotypic and genotypic characteristics	Classical viruses		Giant viruses		Cellular micro-organisms	
	Majority case	Exceptions/ comments	Majority case	Exceptions/ comments	Majority case	Exceptions/comments
Visible under a light microscope (>0.2 μm)	No	–	Yes	–	Yes	–
Genome size > 350 kbp	No		Yes	–	Yes	–
Presence of a virally-encoded capsid	Yes	Some capsidless viruses: genus *Mitovirus*, *Umbravirus*, *Hypovirus*, *Endornavirus* (Koonin and Dolja, 2014)	Yes	Pandoraviruses (Abergel et al., 2015)	No	Icosahedral compartments exist in bacteria and archaea that resemble to viral capsids: the encapsulin nanocompartments structurally similar to and possibly derived from major capsid proteins of tailed bacterial and archaeal caudaviruses, and bacterial microcompartments present in bacteria (including cyanobacteria and many chemotropic bacteria) that encapsulate enzymes involved in metabolic pathways (Tanaka et al., 2008; Krupovic and Koonin, 2017)
Presence of DNA and RNA inside the viral particle	No	Cytomegalovirus (Terhune et al., 2004)	Yes	–	Yes	–
Absolute parasitism	Yes	–	Yes	–	Several bacteria and archeae	Case of strictly intracellular microorganisms
Multiplication by binary fission	No	–	No	–	Yes	No *bona fide* binary fission for *Chlamydia* spp. (Abdelrahman et al., 2016; Bou Khalil et al., 2016), *Ehrlichia* spp. (Zhang et al., 2007), and *Babela* sp. (Pagnier et al., 2015)
Eclipse period during the replicative cycle	Yes	–	Yes	–	No	–
Entry into host cells by phagocytosis	No	–	Yes	–	–	–
Presence of a virus factory	In several viruses (e.g., adenoviruses, polyadenoviruses) (Netherton and Wileman, 2011)	–	Yes	Mollivirus (Legendre et al., 2015)	–	Morula similar to a viral factory for *Chlamydia* spp. (Abdelrahman et al., 2016; Bou Khalil et al., 2016), *Ehrlichia* spp. (Zhang et al., 2007), and *Babela* sp. (Pagnier et al., 2015)
Energy (ATP) generating machinery	No	–	No	–	Yes	Through glycolysis in *Mycoplasma genitalium* (Glass et al., 2006); absence in *Carsonella ruddii* (Nakabachi et al., 2006)

(Continued)

TABLE 1 | Continued

Phenotypic and genotypic characteristics	Classical viruses		Giant viruses		Cellular micro-organisms	
	Majority case	Exceptions/comments	Majority case	Exceptions/comments	Majority case	Exceptions/comments
Presence of genes encoding ribosomal RNA and proteins	No	–	No	–	Yes	Uncomplete sets of ribosomal proteins and aminoacyl-tRNA synthetase in *Carsonella ruddii* (Tamames et al., 2007)
Presence of genes encoding translation-associated proteins	No	–	Yes	–	Yes	–
Presence of tRNA genes	No	–	Yes	Marseilleviruses, faustoviruses (Reteno et al., 2015)	Yes	–
Presence of viral proteins of transcription	Yes	–	Yes	Not detected by proteomics in a marseillevirus (Fabre et al., 2017)	Yes	–
Presence of host ribosomal proteins inside virions	No	Arenaviruses (Baird et al., 2012)	In Mollivirus (Legendre et al., 2015)	–	Yes	–
Presence of group I, II or spliceosomal introns, inteins	No	–	Yes	–	Yes	–
Transposable elements	No	–	Yes	Introns, inteins, transpovirons, miniature inverted-repeat transposable elements (MITEs, in pandoraviruses)	Yes	–
Infection by other viruses	No	–	No	Mimiviruses with (pro)virophages (La Scola et al., 2008; Fischer and Suttle, 2011)		
Mechanism of defense against viruses	No	–	Yes for mimiviruses	–	Yes	–
High level of genome mosaicism	No	–	Yes	–	Yes	–
Evidence of ancestrality based on conserved/ubiquitous genes and protein fold-superfamilies	Four monophyletic classes of viruses (Koonin et al., 2006)	–	Yes	–	Yes	–

are the most conserved between cellular organisms and viruses (Boyer et al., 2010). They notably include three genes conserved among previously identified *Megavirales* representatives and in faustoviruses, and that encode DNA-dependent RNA polymerase subunits 1 (RNAP1) and 2 (RNAP2), and family B DNA polymerase (DNApol). Viral orthologs for these three genes were retrieved with the OrthoMCL program (Li et al., 2003) from the gene complements of 317 viral genomes harboring > 100 genes downloaded from the NCBI sequence databases[1], and orthologs from nine faustovirus genomes (Benamar et al., 2016) and Mollivirus sibericum (Legendre et al., 2015) were added to this sequence set (**Supplementary Table S1**).

Collection of Orthologous Sequences From Cellular Organisms

Informational gene homologs from cellular organisms (maximum number: 20,000) were retrieved from the NCBI GenBank non-redundant (nr) protein sequence database by stand-alone BLAST searches with viral sequences as query, using default parameters except for the maximum target number limit, set to 20,000 (Altschul et al., 1990). Homologous sequences were selected from representative species that diverged approximately 500 million years ago using TimeTree (Hedges et al., 2006; Sharma et al., 2014). BLASTp results were filtered by taxon identifiers, selected sequences were downloaded using their GenBank identifier, and duplicates were removed by clustering with the CD-HIT suite, as previously described (Sharma et al., 2014, 2015b).

Multiple Sequence Alignments and Phylogeny Reconstructions

Sequences (**Supplementary Table S2**) were aligned with the MUSCLE software (Edgar, 2004) and alignments were manually curated. Phylogeny reconstructions were performed using FastTree (Price et al., 2010) with the Maximum Likelihood method, and the CAT 20 model that analyses the alignment site by site and reduces long branch attraction artifacts (Lartillot et al., 2007). Then, trees were visualized using FigTree[2]. Confidence values were determined by the Shimodaira-Hasegawa (SH) test using FastTree (Price et al., 2010).

Comparison of Informational Genes Repertoires

Hierarchical clustering was performed with the Pearson distance method and the TM4 multi-package software, as previously described (Sharma et al., 2015a,b). This analysis relied on the comparison of the presence/absence patterns of 726 COGs involved in nucleotide transport and metabolism and information storage and processing in the gene contents of viruses and of selected bacterial, archaeal, and eukaryotic representatives (Sharma et al., 2015a,b). Viral orthologs were identified through BLASTp searches using these 726 COGs. BLAST searches were performed with default parameters, except for the maximum target number limit, set to 20,000.

Comparison of Gene Repertoires From a Representative of Each of the Three Cellular Domains of Life and From a Giant Virus, and Construction of the Rhizome of Genomes and Genes

Comparison of the gene contents was performed for three members of cellular domains that were selected because they harbor small genomes and are intracellular parasites [namely *Encephalitozoon intestinalis* (an eukaryote) (Corradi et al., 2010), *Methanomassiliicoccus luminyensis* (an archaeon) (Gorlas et al., 2012), *Rickettsia bellii* (a bacterium) (Ogata et al., 2006)], and for Tupanvirus soda lake (Abrahao et al., 2018), a recently described giant virus that was selected here because it has a particularly large gene content and harbors the largest set of translation components among giant viruses. This comparison used the ProteinOrtho v5 tool with 1e-3, 20 and 30% as thresholds for e-value, amino acid identity, and coverage of aligned sequences, respectively (Lechner et al., 2011). In addition, best BLASTp hits against the NCBI GenBank protein sequence database were obtained for these four organisms. The "rhizomes" of the genomes were built using the Circos tool[3]. Rhizomes consist in a representation of the genome evolution and mosaicism that takes into consideration the fact that genes from this genome as well as intragenic sequences do not have the same evolutionary history, and can result from exchanges, fusions, recombination, degradation, or *de novo* creation (Raoult, 2010). Rhizomes, which are devoid of a center, were proposed as a better paradigm of genetic evolution than trees (Deleuze and Guattari, 1976; Raoult, 2010). Rhizomes built here show in a single figure, for all the genes from a given virus or cellular organisms, the taxonomy of their best BLASTp hits that represent putative donors or acceptors involved in sequence transfers, as well as the ORFans (sequences devoid of homolog in databases). Furthermore, a rhizome of genes was also determined for the genes encoding a methionyl-tRNA synthetase shared by the four organisms, by performing BLASTp searches with fragments obtained from this gene by cutting its amino acid sequence into 40 amino acid-long fragments that overlapped with a sliding window of 20 amino acids.

RESULTS AND DISCUSSION

Phylogenetic Analyses of Protein Structural Domains of Viral and Cellular Proteomes

A total of ∼1,200 folds, ∼2,000 superfamilies, and ∼5,000 families of structural domains encompass the entire evolutionary and functional diversity of the protein world. The history of these folds, superfamilies and families has been traced with phylogenomic methods by studying the entire repertoires of

[1] ftp://ftp.ncbi.nih.gov/genomes/Viruses/

[2] http://tree.bio.ed.ac.uk/software/figtree/

[3] http://circos.ca/

proteins (proteomes), beginning with a study of a small set of 32 completely sequenced genomes (Caetano-Anollés and Caetano-Anollés, 2003) and continuing with a recent extended analysis of thousands of viral and cellular genomes (Nasir and Caetano-Anollés, 2015). Timelines of domain history could be calibrated with a molecular clock that relates them to the geological record (Wang et al., 2011). The timelines showed that the oldest domain families harbored 'Rossmann-like' α/β/α-layered and bundle structures typical of globular proteins, followed by barrel structures typical of membrane and metabolic proteins (Caetano-Anollés et al., 2012). The oldest of these structures are predominant in membrane-associated proteins, suggesting a very early onset of cellular structure. Their link to metabolism, but not translation, also suggests the late development of the genetic code and the late appearance of the ribosome (Harish and Caetano-Anollés, 2012; Caetano-Anollés et al., 2013). Remarkably, the late arrival of modern genetics ∼3 billion years (Gy) ago signals the end of a period responsible for the primordial cellular origin of viruses, clearly evident by the fact that the oldest superfamilies are common to cells and viruses (Nasir and Caetano-Anollés, 2015). In addition, these data also indicated that RNA polymerases are more ancient than the ribosome. Such diversification occurred prior to the appearance of the cellular domains of life.

A previous phylogenomic data-driven analysis of proteomes confirmed the early cellular origin of viruses and the rise of viral RNA proteomes followed by that of DNA viruses and Megavirales representatives (Nasir and Caetano-Anollés, 2015). Here we focused on the evolutionary relationship of Megavirales and cellular organisms. Out of all possible FSF domains (**Figure 1**), we selected 289 that were universal, i.e., that were shared by viruses and cellular organisms. We then used this set to build a phylogeny of proteomes (**Figure 2**). Megavirales representatives appear as a basal group in the tree of proteomes, which is consistent with results from sequence analyses performed here and previously (Boyer et al., 2010; Sharma et al., 2014). The subgroup that was closest to cellular organisms was family *Mimiviridae*, followed by family *Phycodnaviridae* and then groups comprised by family *Marseilleviridae* and by faustoviruses, mollivirus, and pandoraviruses. Similar phylogenetic patterns were revealed when we used multidimensional scaling approaches to explore the temporal space of ages of individual structural domains in proteomes (**Figure 3**). We found distinct temporal clouds of proteomes for viruses and organisms belonging to *Archaea*, *Bacteria,* and *Eukarya*. The *Mimiviridae* group was clearly dissected from the main viral cloud, which was temporally closer to cellular proteomes, suggesting their late appearance in viral evolution. Again, the family *Phycodnaviridae* appeared between the family *Mimiviridae* and the rest of the viral cloud. In terms of the proportions of FSFs detected in giant viral groups, asfarviruses have a proteome that is more similar to that of faustovirus, which is consistent with phylogenetic analysis of sequences. However, when considering raw number, mimiviruses have more FSFs in common with faustovirus. Finally, when plotting phylogenetic indices measuring the levels of homoplasy of the MP tree reconstruction (corresponding to **Figure 2**) against age of the phylogenetic character (fold superfamily), high retention indices, especially for lower *nd* values (oldest domains), indicated excellent fit of characters to the phylogeny (**Figure 4**). Homoplasy indicates the level of independent gain of characters in lineages and is a good indicator of deviations from vertical inheritance (Farris, 1983). The levels of homoplasy were moderate for protein folds, showing that the vertical signals override the horizontal signals.

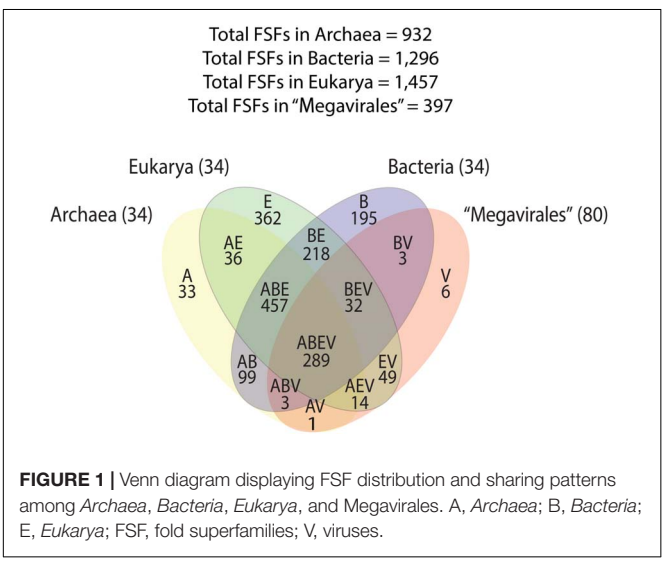


FIGURE 1 | Venn diagram displaying FSF distribution and sharing patterns among *Archaea*, *Bacteria*, *Eukarya*, and Megavirales. A, *Archaea*; B, *Bacteria*; E, *Eukarya*; FSF, fold superfamilies; V, viruses.

Phylogenetic Analyses of RNA and DNA Polymerases and Phenetic Comparison of Informational COGs

As shown in **Figures 5**, **6**, trees reconstructed using both RNA polymerase subunit sequences (RNAP1 and 2) from members of Megavirales (including recently described giant viruses of amoebae), *Bacteria*, *Archaea,* and *Eukarya* clearly displayed a topology with four branches. The Megavirales group exhibits a considerable genetic diversity. Regarding phylogeny reconstruction based on DNA polymerases present in archaea, eukaryotes and giant viruses, giant viruses are separated into two groups. Faustoviruses and asfarviruses are clustered together and comprise sister branches, apart from other giant viruses that form an independent and strongly supported cluster (**Figure 7**). Hierarchical clustering analysis was performed based on a binary presence/absence matrix constructed using 727 informational COGs present in 143 representative genomes of cellular organisms from *Bacteria*, *Archaea* and *Eukarya*, and viruses from Megavirales (**Figure 8**). This phenetic analysis based on informational genes also showed a four-branch topology, Megavirales being a distinct branch alongside *Eukarya*, *Archaea,* and *Bacteria*.

Pangenome and Core Genome for One Member of Each of the Three Cellular Domains of Life and of a Giant Virus

A pangenome and core genome was determined for one representative of each of the four TRUCs of microbes: namely

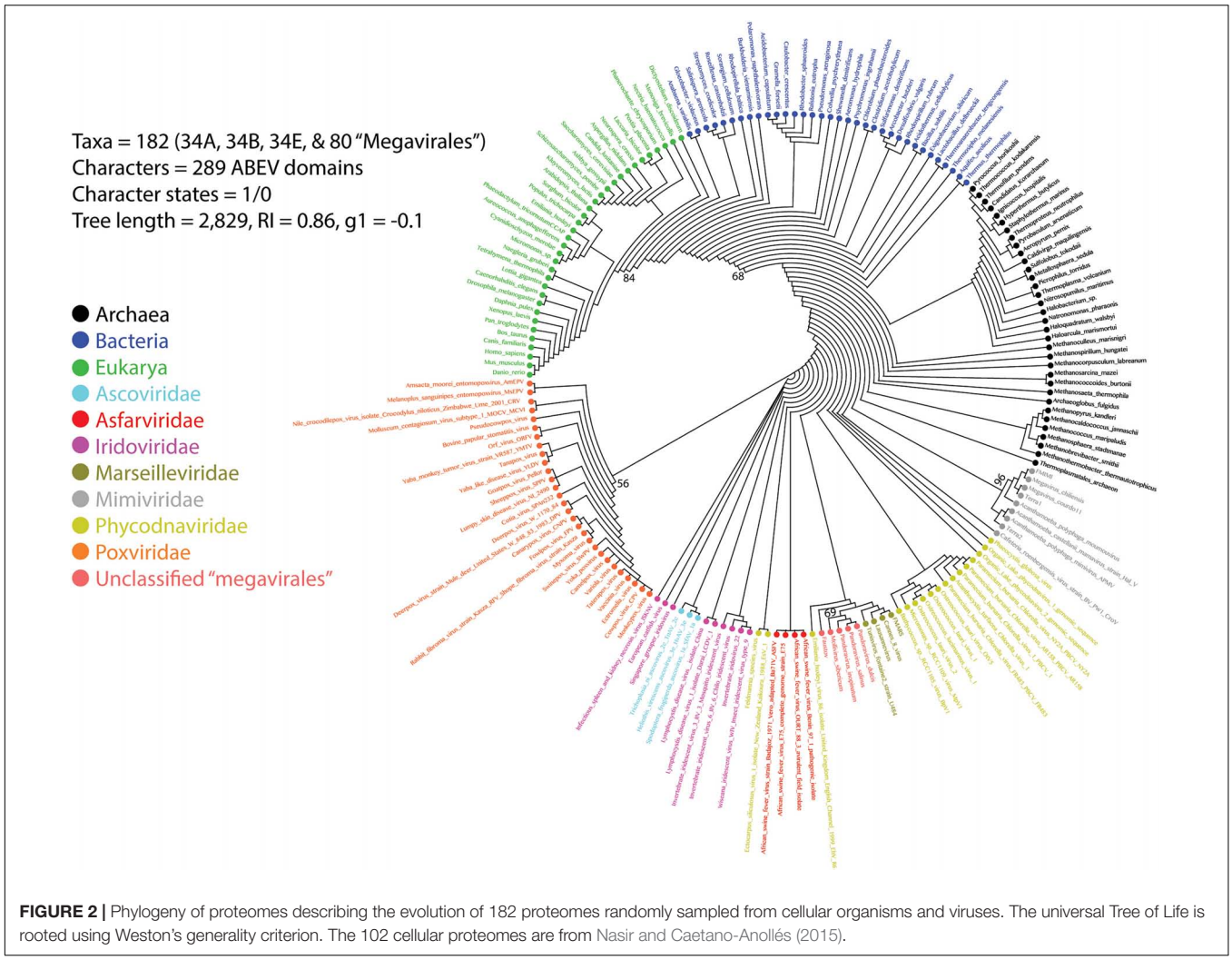


FIGURE 2 | Phylogeny of proteomes describing the evolution of 182 proteomes randomly sampled from cellular organisms and viruses. The universal Tree of Life is rooted using Weston's generality criterion. The 102 cellular proteomes are from Nasir and Caetano-Anollés (2015).

R. bellii (bacteria, 1,430 genes), *M. luminyensis* (archaea, 2,533 genes), *E. intestinalis* (eukaryota, 1,910 genes), and Tupanvirus soda lake (giant virus, 1,269 genes). The pangenome describes the full complement of genes in a group of organisms, in our case the four microbes, and is comprised by the core genome that contains genes present in all 4 microbes and by the dispensable genome composed of genes that are unique to each microbe and genes absent from one or more microbes. The pangenome of these four microbes was composed of 6,531 genes, and their core genome (shared by all four organisms) was composed of 33 genes that represented between 1.3 and 2.6% of their gene contents. This core genome included notably genes encoding a DNA-directed RNA polymerase, a ribonucleoside-diphosphate reductase, a translation elongation factor 2, and several aminoacyl-tRNA synthetases. A majority of these genes therefore consisted of translation components. In addition, 23 (1.6%), 68 (5.4%), 13 (0.7%), and 68 (5.4%) genes from *R. bellii*, *M. luminyensis*, *E. intestinalis*, and Tupanvirus, respectively, had homologs in the genomes of two other microbes. Finally, 261 genes in *R. bellii* (18.3%), 362 in *M. luminyensis* (14.3%), 298 in *E. intestinalis* (15.6%), and 132 in Tupanvirus (10.4%) had homologs in at least one of the three other microbes. These results show that beyond the fact that the number of genes for Tupanvirus is in the same order of magnitude than for the three cellular microorganisms, a substantial proportion of the genes of this giant virus overlaps with those of the bacteria, the archaeon and the eukaryote.

Rhizomes of Genomes and Genes as Appropriate Representations of the Origin and Evolution of Members From the Four TRUCs of Microbes

A substantial genome mosaicism, consisting of genomes composed by genes with sequences suggesting different evolutionary origins and histories, was observed for representatives of the four TRUCs, including *R. bellii*, *M. luminyensis*, *E. intestinalis*, and Tupanvirus (**Figure 9**). This mosaicism was particularly predominant in the Tupanvirus genome as described previously (Abrahao et al., 2018), with 51, 11, 8, and 0.2% of its genes best matching with viruses,

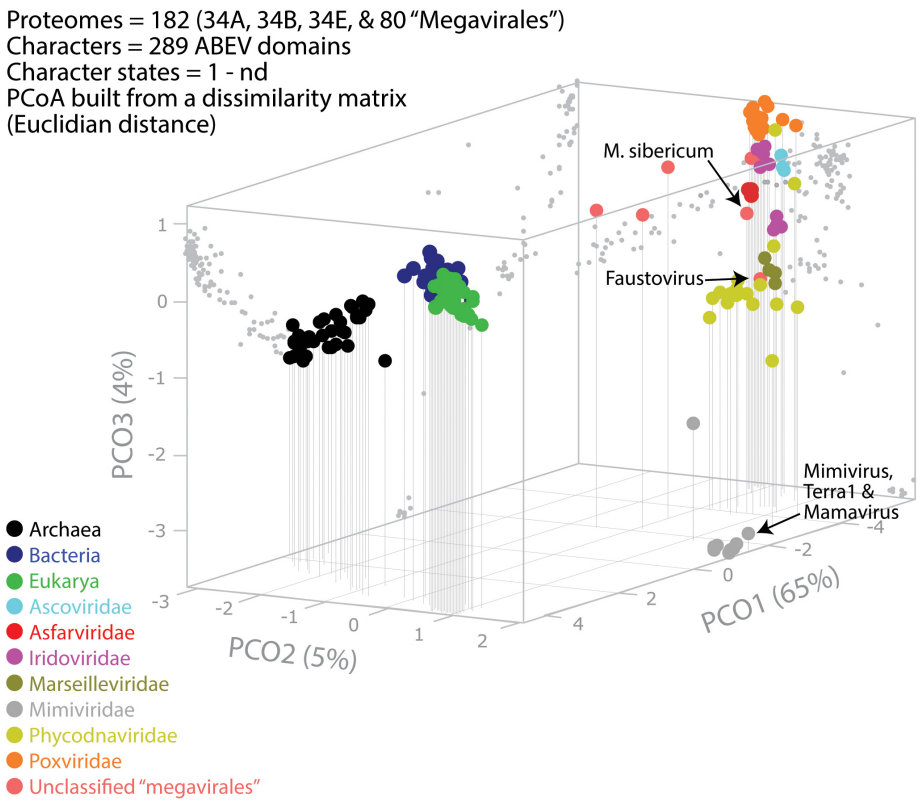


FIGURE 3 | Evolutionary principal coordinate (evoPCO) analysis plot portrays in its first three axes the evolutionary distances between cellular and viral proteomes. The percentage of variability explained by each coordinate is given in parentheses on each axis. Data points of the 3-dimensional scatter plot describing temporal clouds are mapped onto projections planes and connected with vertical leading drop lines along the PCO3 axis. The list of whole coordinate information for building the PCoA plot of this figure is provided in **Supplementary Table S3**.

eukaryota, bacteria, and archaea, respectively, but it was a shared feature of the three non-eukaryotic microorganisms. This illustrates that a rhizome is the most appropriate representation of the evolutionary history at a genome scale, as individual genes can have distinct and distant origins (Raoult, 2010). Such representation notably takes into account introgressive descent as a result of lateral sequence transfers. Moreover, it appears that genes themselves may be subject to lateral sequence transfer rearrangements (through gene conversion), as shown here for the case of the methionyl-tRNA synthetase encoding gene of the four microorganisms (**Figure 10**). Indeed, 40 amino acid-long fragments of these genes alternately found as best hits, apart from relatives from the same family or genus, sequences from archaea, bacteria, eukaryota, or viruses. Such a gene sequence mosaicism was particularly broad for Tupanvirus and *M. luminyensis*. For the case of Tupanvirus soda lake, 15, 3, 2, and 1 methionyl-tRNA synthetase gene fragments found as best hits an eukaryote, a virus, a bacterium and an archaeon, respectively. This was also remarkably exemplified with the case of the glutaminyl-tRNA synthetase of Klosneuvirus, a mimivirus relative (Schulz et al., 2017). Indeed, fragments of this glutaminyl-tRNA synthetase gene showed a mixture of sequences from eukaryotes, bacteria and of unknown sources, or of sequences retrieved from metagenomes, in particular those of Antarctic dry valleys (Abrahao et al., 2018). These findings make the notion of gene lateral transfer obsolete, as sequences, rather than genes, are transferred (Merhej et al., 2011). Thus, the source of a gene may be better defined by a rhizome than by a tree, as previously proposed for organisms (Raoult, 2010) (**Figure 11**). Examples of chimeric genes have been previously described. Thus, ORF13 of the Sputnik virophage encodes a primase-helicase whose N-terminal region is of archaea-eukaryotic source and C-terminal portion was inferred to originate from giant viruses (La Scola et al., 2008). In the fern *Adiantum capillus-veneris*, a chimeric photoreceptor was identified that may have been critical in the divergence and rise of some fern species under low luminosity environments (Kawai et al., 2003). More broadly, it has been described that the creation of novel chimeric genes, referred as chimeric nuclear symbiogenetic genes (S-genes), occurred during eukaryogenesis through the fusion of bacterial and archaeal genes; this gave rise in early eukaryotes to novel chimeric proteins with central functions (Meheust et al., 2018). These data confirm and expand to genes the concept that no single tree can define the chimeric nature of genomes, as genes themselves are mosaics (Dagan and Martin, 2006; Merhej et al., 2011). As a consequence, trees made with homologous sequences make no sense if

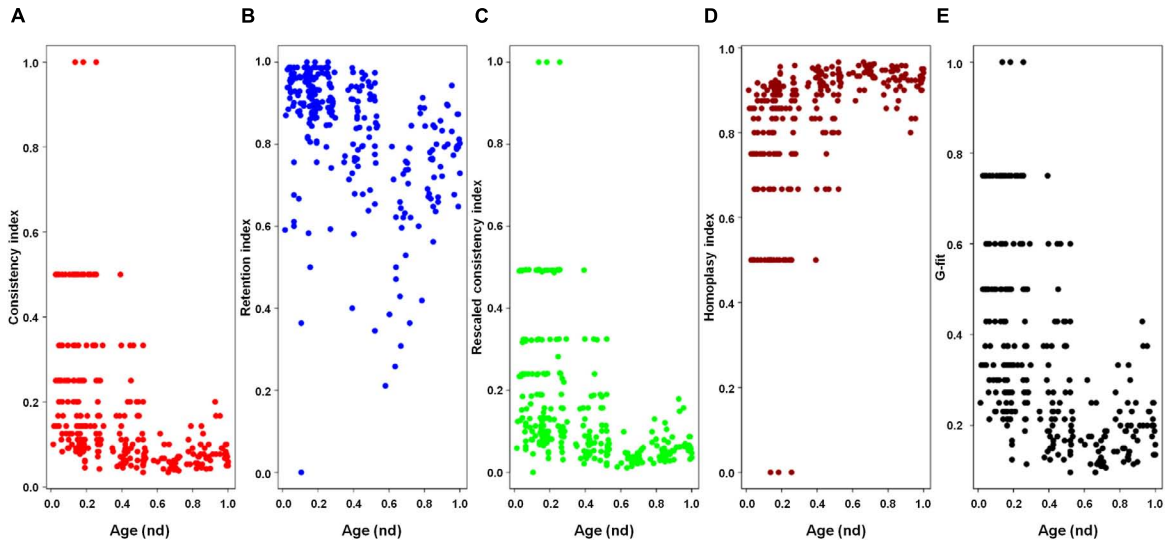


FIGURE 4 | Plots of the indices of the phylogenetic tree of proteomes describing the evolution of 182 proteomes randomly sampled from cellular organisms and viruses (corresponding to **Figure 2**) against the age of the phylogenetic character [fold superfamily (FSF)]. Five measures of the levels of lateral sequence transfers for the maximum parsimony tree reconstruction performed in the present study, namely consistency index **(A)**, retention index **(B)**, rescaled consistency index **(C)**, homoplasy index **(D)**, and G-fit **(E)**, are plotted against the age of the phylogenetic character FSF [measured as node distance (*nd*) values] for 289 characters (FSF) shared by archaea, bacteria, eukaryota, and viruses. High retention indices, especially for lower *nd* values (corresponding to older domains), indicates excellent fit of the characters to the phylogeny.

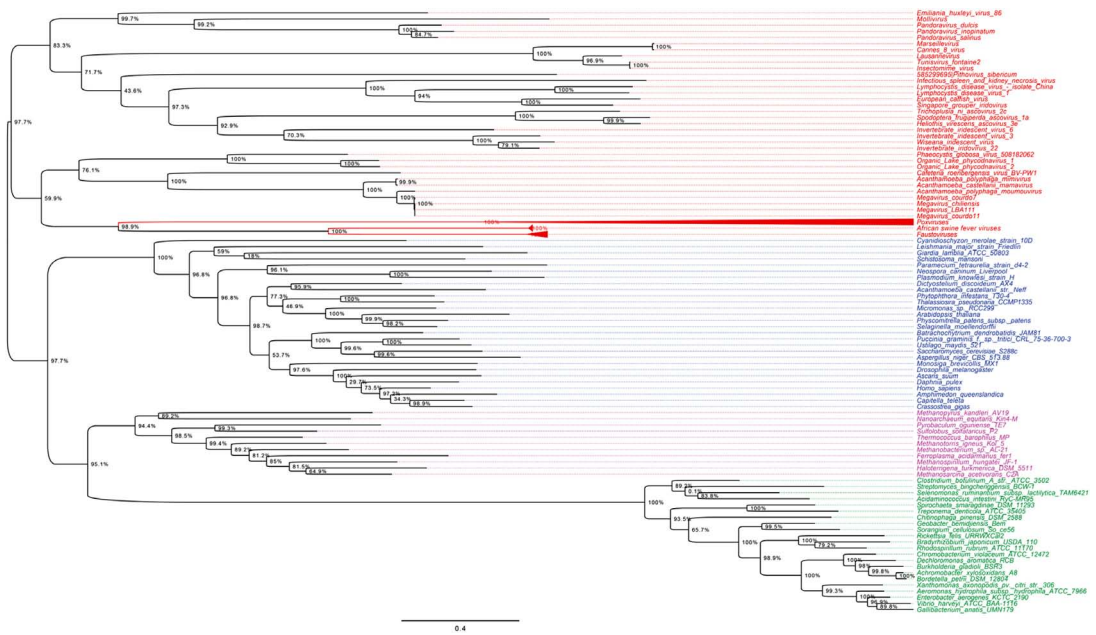


FIGURE 5 | RNAP1 phylogenetic tree. The RNAP1 tree was built by using aligned protein sequences from *Megavirales* (red), *Bacteria* (green), *Archaea* (pink), and *Eukarya* (blue). Confidence values were calculated by the Shimodaira-Hasegawa (SH) test using the FastTree program (Price et al., 2010). Average length of sequences was 1,336 amino acids. The scale bar represents the number of estimated changes per position.

not all fragments of these sequences have a common source. Phylogeny reconstructions based on concatenated genes are still worse when the trees built based on the separate genes do not have the same topology, because they consist in mixing sequences from different, and eventually very distant, origins.

Definition Criteria for Giant Viruses or Megavirales

As shown in **Table 1**, giant viruses exhibit unique phenotypic and genotypic features that differentiate them from 'classical' viruses, indicate their much greater complexity, and bring them close

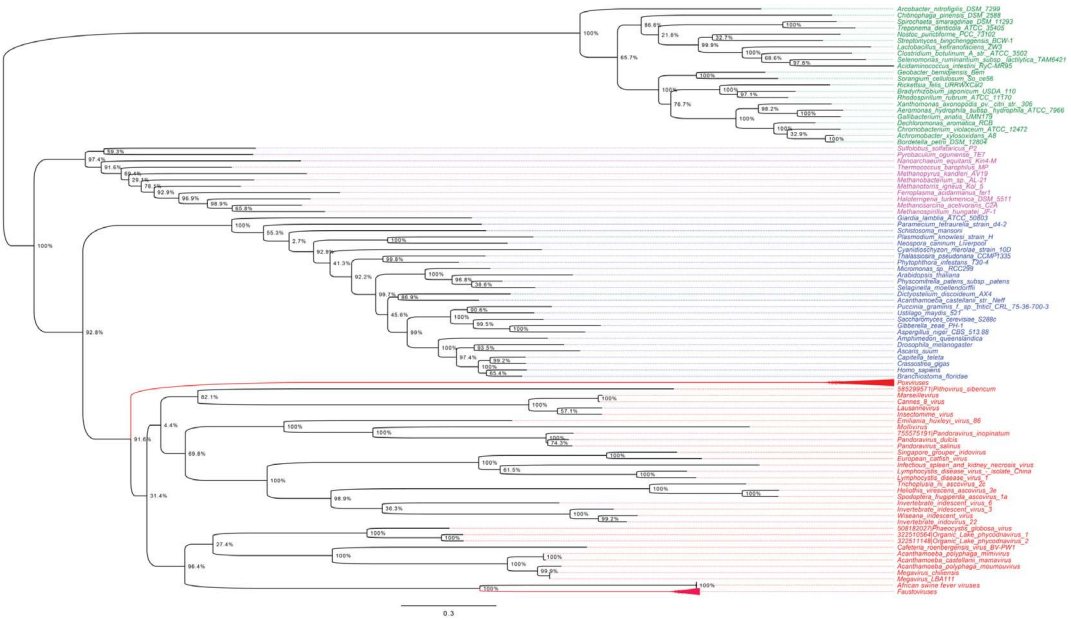


FIGURE 6 | RNAP2 phylogenetic tree. The RNAP2 tree was built by using aligned protein sequences from *Megavirales* (red), *Bacteria* (green), *Archaea* (pink), and *Eukarya* (blue). Confidence values were calculated by the SH test using the FastTree program (Price et al., 2010). Average length of sequences was 1,188 amino acids. The scale bar represents the number of estimated changes per position.

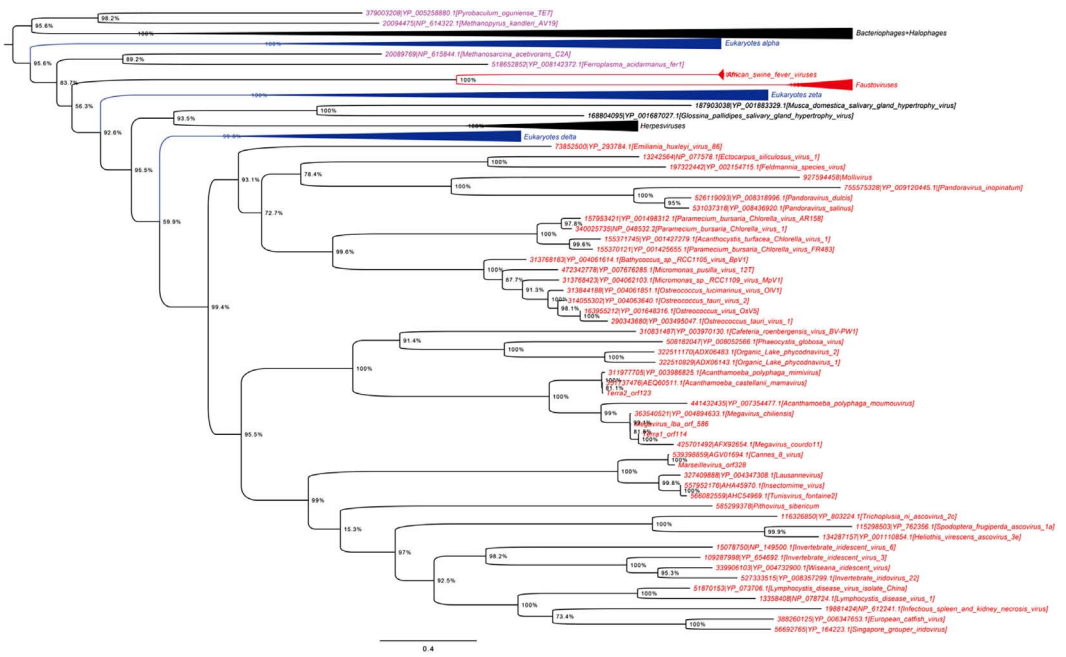


FIGURE 7 | DNA polymerase phylogenetic tree. The DNA polymerase tree was built by using aligned protein sequences from *Megavirales* (red), *Bacteria* (green), *Archaea* (pink), and *Eukarya* (blue). Confidence values were calculated by the SH support using the FastTree program (Price et al., 2010). Average length of sequences was 1,134 amino acids. The scale bar represents the number of estimated changes per position.

to small micro-organisms. These characteristics can be classified as follows: (i) Giant sizes of the virions and their genomes. (ii) Complexity, with presence in virions of dozens of proteins, and of messenger RNA. (iii) Presence of translation components unique among viruses; in this view, the recent characterization of klosneuviruses (Schulz et al., 2017) and tupanviruses (Abrahao et al., 2018) has led to a considerable expansion of the set of such translation components. Notably, the tupanvirus isolates

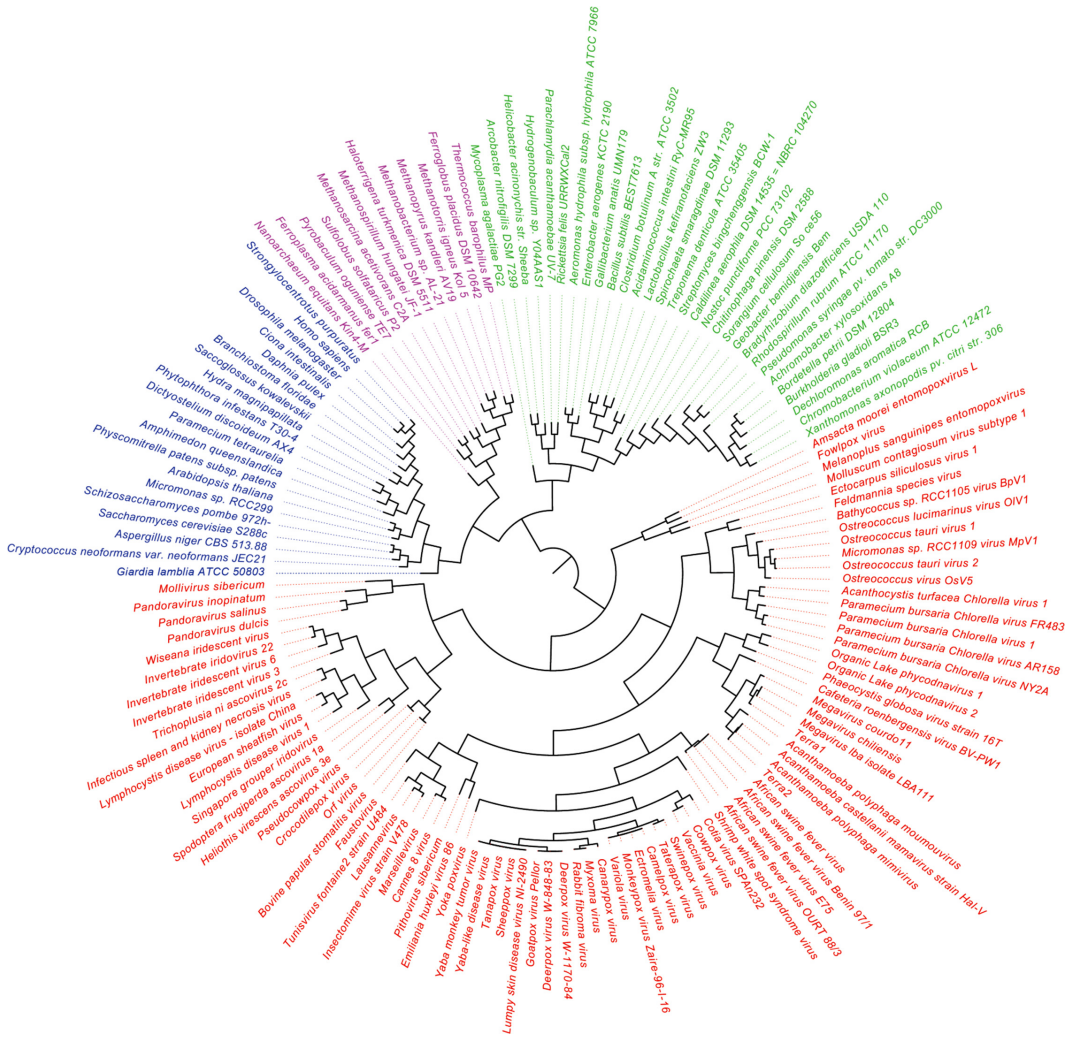


FIGURE 8 | Hierarchical clustering by phyletic pattern based on the presence/absence of informational Clusters of Orthologous Groups (COGs) of proteins. The Megavirales members are represented in red, *Bacteria* members in green, *Archaea* members in pink, and *Eukarya* members in blue.

encode for 67–70 tRNA, 20 aminoacyl tRNA-synthetases, and 11 translation factors. (iv) Presence of a specific mobilome in mimiviruses that includes virophages, transpovirons, introns, and endonucleases (Desnues et al., 2012), as well as MIMIVIRE, a defense system against virophages (Levasseur et al., 2016b; Dou et al., 2018). (v) Based on phylogenetic, phyletic, and protein fold superfamilies analyses, delineation of a fourth group of micro-organisms comprised by giant amoebal viruses alongside bacterial, archeal and eukaryotic microbes, and evidence of an archaic origin (Boyer et al., 2010; Sharma et al., 2014; Nasir and Caetano-Anollés, 2015). Moreover, the recent comparison of the genomes of a fossil and a modern pithovirus highlighted that giant viruses evolve with a mutation rate estimated to be lower than that of RNA viruses and comparable to those determined for bacteria and archaea, and by classical mechanisms of evolution, including through long-term fixation of genes that are acquired by horizontal gene transfer (Levasseur et al., 2016a).

Giant viruses of amoebae certainly exhibit several criteria that are hallmarks and definition criteria of viruses. These include the occurrence of an eclipse phase during their replicative cycle, an obligatory replication into host cells, and the presence of a capsid (Lwoff, 1957; La Scola et al., 2003; Raoult and Forterre, 2008). Nevertheless, regarding the capsid, pandoraviruses, pithoviruses, mollivirus, and cedratviruses have virions surrounded by a tegument-like structure and no known capsid morphology (Philippe et al., 2013; Yutin and Koonin, 2013; Legendre et al., 2014, 2015). Pandoraviruses do not have a recognizable capsid-encoding gene, pithoviruses have a barely identifiable capsid-encoding gene, while capsid proteins are detected in Mollivirus virions but they are not part of the virion structure. Other giant virions with an ovoid or spherical shape such as cedratviruses and Orpheovirus are also devoid of a morphology resembling those provided by known capsids. An atypical capsid structure was previously described for Megavirales representatives. Thus, most poxviruses have

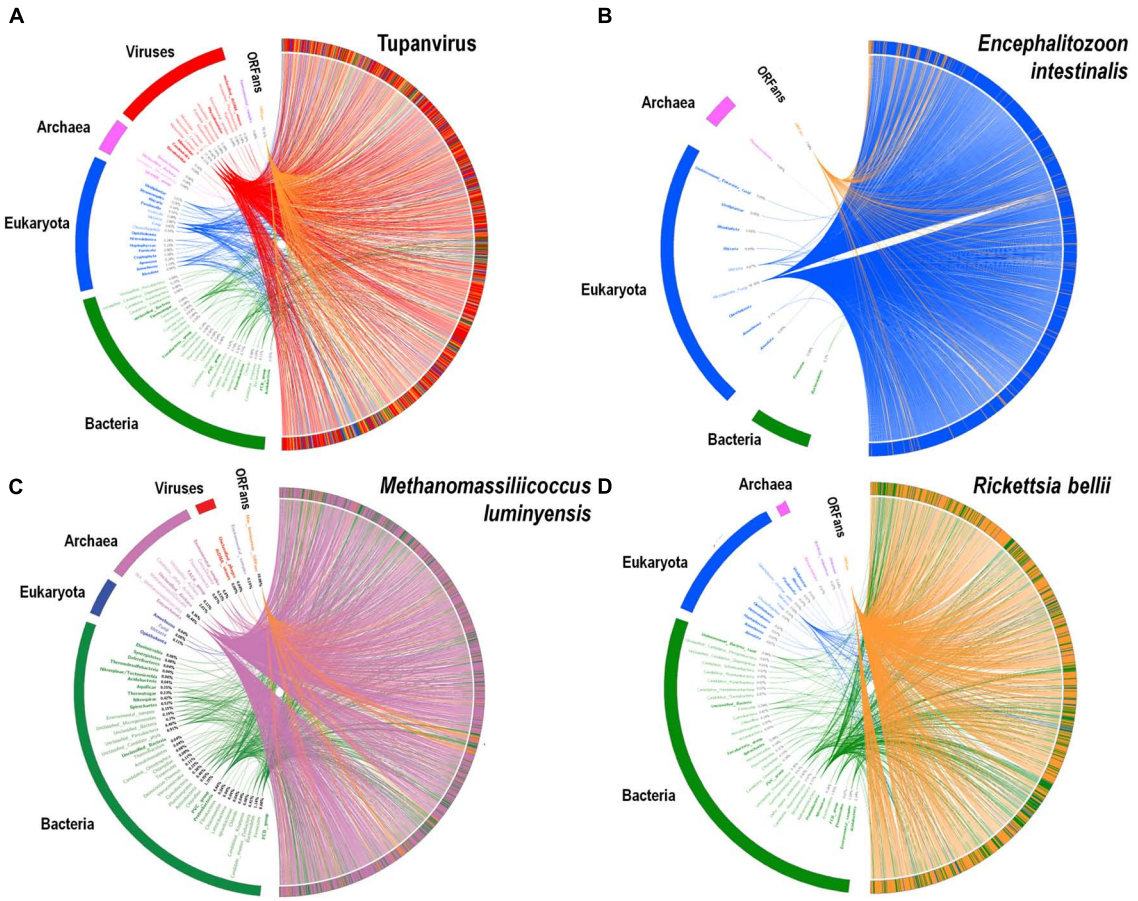


FIGURE 9 | Rhizomes of genomes illustrative of the mosaicism of the genomes of representatives of the four TRUCs of microbes including Tupanvirus soda lake (a mimivirus) **(A)**; *Encephalitozoon intestinalis* (a microbial eukaryote) **(B)**; *Methanomassiliicoccus luminyensis* (an archaeon) **(C)**; and *Rickettsia bellii* (a bacterium) **(D)**. The genes of these four microorganisms were linked to their most similar sequences in the NCBI GenBank protein sequence database according to the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi), classified according to their belonging to viruses, eukaryotes, bacteria or archaea, and integrated in a circular gene data visualization. The figures were performed using the CIRCOS online tool (http://mkweb.bcgsc.ca/tableviewer/visualize/). Circular representations in **A** and **C** are the same than those produced for figures from articles Abrahao et al. (2018) and Levasseur et al. (2017), respectively, as they originate from the same data. These representations are licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/) and CC-BY-NC (https://creativecommons.org/licenses/by-nc/4.0/), respectively.

brick-shaped virions, the capsid precursors being assembled following icosahedral symmetry and the final shape being reached after proteolytical cleavages (Condit et al., 2006), and ascoviruses harbor allantoid capsids (Federici et al., 1990).

Moreover, although giant viruses of amoebae share phenotypic and genotypic features with cellular microorganisms, they were described to lack key cellular hallmarks. A first one consists in proteins involved in the production of energy. This might not be strictly true as tupanviruses harbor genes encoding a putative citrate synthase (Abrahao et al., 2018), and the genome of a distant mimivirus relative (Tetraselmis virus 1) that infects a green alga was shown to harbor key fermentation genes (a pyruvate formate-lyase and a pyruvate formate-lyase activating enzyme) that might ensure energy requirements (Schvarcz and Steward, 2018). A second one consists in ribosomal DNA and proteins, which are absent from giant viruses. Nevertheless, two distinct copies of an 18S rRNA intronic region were recently described in tupanviruses (Abrahao et al., 2018). These sequences were found to be highly expressed, and led to detect similar 18S rRNA intronic region in the majority of other mimivirus genomes. A third cellular hallmark that lacks in giant viruses of amoebae is binary fission as multiplication mechanism.

Conversely, it must be also considered that some bacteria display viral specific features and also lack hallmark features of cellular microorganisms. Numerous bacteria are indeed obligatory intracellular parasites. Moreover, some small cellular microorganisms such as *Carsonella ruddii* lack a comprehensive ATP generation machinery and, in addition, have a not comprehensive set of ribosomal proteins and aminoacyl-tRNA synthetases (Nakabachi et al., 2006; Tamames et al., 2007). Other cellular microorganisms, such as *Chlamydia* spp. (Abdelrahman et al., 2016; Bou Khalil et al., 2016), *Ehrlichia* spp. (Zhang et al., 2007), and *Babela* sp. (Pagnier et al., 2015) have no *bona fide* binary fission step during their multiplication. These data

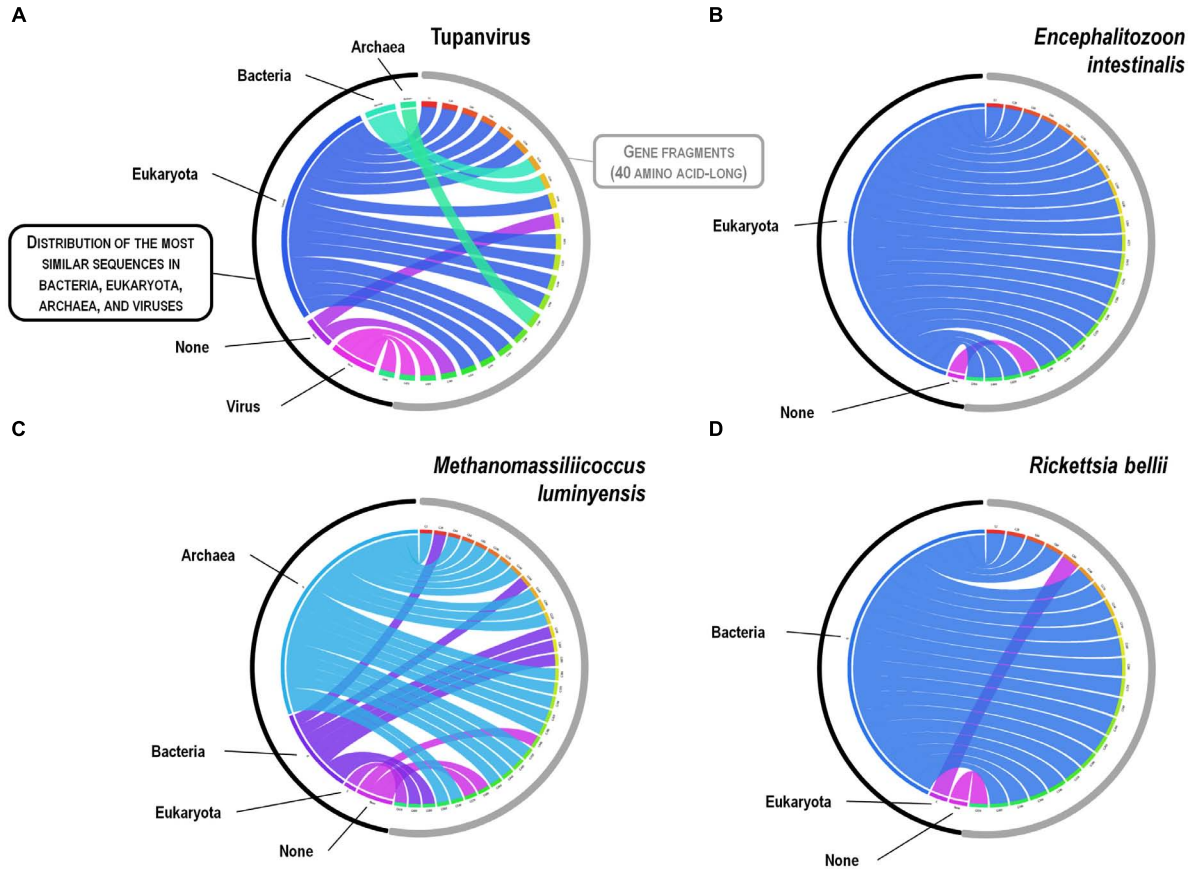


FIGURE 10 | Rhizomes of methionyl-tRNA synthetase gene fragments illustrative of the mosaicism of the genes of representatives of the four TRUCs of microbes including Tupanvirus soda lake (a mimivirus) **(A)**; *Encephalitozoon intestinalis* (a microbial eukaryote) **(B)**; *Methanomassiliicoccus luminyensis* (an archaeon) **(C)**; and *Rickettsia bellii* (a bacterium) **(D)**. Forty amino acid-long fragments of the methionyl-tRNA synthetase encoding genes of the four microorganisms were linked to their most similar sequences in the NCBI GenBank protein sequence database according to the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi), classified according to their belonging to viruses, eukaryotes, bacteria or archaea, and integrated in a circular gene data visualization. The figures were performed using the CIRCOS online tool (http://mkweb.bcgsc.ca/tableviewer/visualize/).

highlight that both classical viruses and cellular microorganisms can lack one or several pillar defining features. Finally, while a few viruses, including pandoraviruses, are devoid of capsid (Philippe et al., 2013; Koonin and Dolja, 2014), two classes of icosahedral compartments exist in bacteria and archaea that resemble to viral capsids: they include encapsulin nanocompartments structurally similar to and possibly derived from major capsid proteins of tailed bacterial and archaeal caudaviruses, and microcompartments present in bacteria (including cyanobacteria and many chemotropic bacteria) that encapsulate enzymes involved in metabolic pathways (Tanaka et al., 2008; Krupovic and Koonin, 2017).

CONCLUSION AND PERSPECTIVES

Viruses have long been considered as parasitic entities invisible by light microscopy and with a limited repertoire of genes (Raoult and Forterre, 2008). The fact that they are devoid of ribosomal genes has confined them outside of the "tree of life." Giant viruses of amoebae have undermined this paradigm due to their characteristics that are, at the scale of classical viruses, outstanding (Raoult et al., 2007; Sharma et al., 2016). Phylogenies that were constructed here based on three ancient genes, including RNAP1/2 and DNA polymerase, delineate a fourth TRUC of microbes, as previously reported (Boyer et al., 2010; Sharma et al., 2014, 2015b). Hierarchical clustering performed using a set of informational COGs also shows a fourth independent branch alongside the three cellular branches. Because the tree of proteomes provides a more global and conserved phylogenomic view of protein domain composition in proteomes, their topologies can differ from single-gene based phylogenies that can independently indicate different evolutionary histories. However, here, the four branch topology was maintained in both sequence and structure based trees.

With the recent expansion of the proposed order Megavirales, the number of genes that are shared by these viruses and cellular organisms has shrunk, making it more difficult to build a fourth branch. Nevertheless, among the genes that still

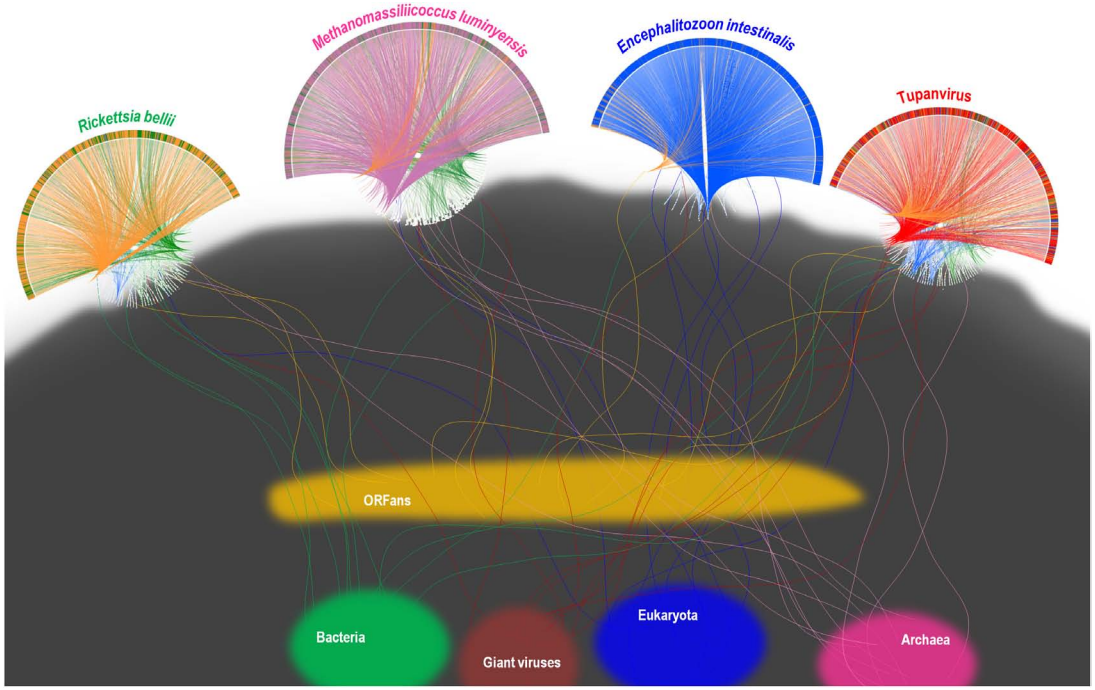


FIGURE 11 | Representation as a rhizome of the genetic evolution for four current intracellular parasites of the four TRUCs of microbes with a comparable genome size, including *Rickettsia bellii* (a bacterium), *Methanomassiliicoccus luminyensis* (an archaeon), *Encephalitozoon intestinalis* (a microbial eukaryote), and Tupanvirus soda lake (a mimivirus). Rhizomes consist in a representation of genome evolution and mosaicism that takes into account that genes and intragenic sequences do not have the same evolutionary history, being proposed as better paradigm of genetic evolution than phylogenetic trees. The genomes of each of the four represented current microorganisms harbor mixtures of sequences of different origins. Sequences corresponding to current bacteria, Archaea, eukaryota, giant viruses, and to ORFans are colored in green, purple, blue, red, and orange, respectively. Rhizomes of the genomes of Tupanvirus and *Methanomassiliicoccus luminyensis* were adapted from same representations than representations from Levasseur et al. (2017) and Abrahao et al. (2018), respectively, licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/) and CC-BY-NC (https://creativecommons.org/licenses/by-nc/4.0/), respectively (see legend to **Figure 9**).

show a monophyly are polymerases, which were shown to be among the most ancient protein fold superfamilies (Nasir and Caetano-Anollés, 2015). The ancestrality of conserved genes such as the RNA polymerases, which are suspected to be more ancient than the ribosome (Nasir and Caetano-Anollés, 2015), highlights that evolution can be the result of structural constraints. This concept was described by Gould and Lewontin who used San Marco Cathedral's spandrels to illustrate that adaptation through selection cannot comprehensively explain the evolution of genomes, and that biological constraints have to be considered (Gould and Lewontin, 1979). The structural, functional and evolutionary units of proteins are the structural domains, highly compact and recurrent segments of the molecules that often combine with others to perform major molecular and cellular tasks (Caetano-Anollés et al., 2009). Domains are evolutionarily highly conserved since they are defined by three-dimensional (3D) structural folds rather than amino acid sequences (Illergard et al., 2009). A rough estimate of evolutionary change suggests that a new fold structure takes millions of years to unfold, while a stable new sequence appears on Earth at least once every microsecond (Caetano-Anollés et al., 2009). In addition, hairpin-forming palindromes, which are possible primordial functional RNAs, are widely distributed among living entities, and they were found to be represented in giant viruses and virophages (Seligmann and Raoult, 2016). Short hairpin structures exist in the genomes of Mimivirus and the Sputnik virophage that may be involved in determining the polyadenylation site of transcripts (Byrne et al., 2009; Claverie and Abergel, 2009). While viral diversification appears fundamentally tailored by reductive evolution, the enrichment of viral genomes with primordial superfamilies of structural domains provides a strong support to the development of the viral proteome core prior to the inception of the ribosome but after the appearance of synthetase-like proteins capable of specific aminoacylation of tRNA molecules (Nasir and Caetano-Anollés, 2015). This could explain the existence of remnants of the translation machinery, the number of which has recently expanded considerably through the isolation of tupanviruses (Abrahao et al., 2018) and the assembly of klosneuvirus genomes (Schulz et al., 2017). As a matter of fact, it is unlikely that there has been a gradual and random acquisition of such large numbers of translation components in giant viruses, such as in mimiviruses, without using it. Hence, this translation machinery might have been acquired in a single step, or, alternatively, might have originated with giant viruses.

The classification of microbes, including the giant viruses, is more realistically based on their genomic content, which reflects their lifestyle, rather than on the phylogenies of supposedly

representative genes, which may be confusing because of their mosaicism. This mosaicism results from sequence (and not gene) exchanges occurring during billion years of interactions between emerging lineages or organisms, and is particularly frequent between sympatric microorganisms (Moliner et al., 2010; Raoult and Boyer, 2010). Indeed, microorganisms that encounter and multiply or replicate in same biological niches are particularly prone to exchange nucleic acid sequences. This is well-suggested by the case of *Acanthamoeba* spp. that can be infected concomitantly by several amoeba-resistant microorganisms including intracellular bacteria and giant viruses with significantly larger repertoires than other related organisms (Moliner et al., 2010; Raoult and Boyer, 2010). Genes evolve by point mutations, but also by fusion, shuffling and fission of genetic fragments, which likely produce gene sequences that are mosaics (Long et al., 1999; Meheust et al., 2018; Pathmanathan et al., 2018). Such chimeric genes have been described in several studies (Ben et al., 2008; Merhej et al., 2011; Meheust et al., 2018), and we found here hints of such gene sequence mosaicism. In addition, many of the genes studied here encode for multi-domain proteins, which makes them mosaics of domains of different ages and histories. The phylogenomic tree reconstructed from domain structures that we describe here disentangles evolutionary histories because each domain becomes a separate phylogenetic character used to build the tree of proteomes. We note however that structural domains and their complex 3D topologies are also built from smaller module-like pieces of arrangements of helix, strand and turn segments (e.g., αα-hairpins, ββ-hairpins, βαβ-motifs) that act as evolutionary building blocks. Recent studies identified combinable (Goncearenco and Berezovsky, 2015) and no-combinable (Alva et al., 2015) 'loop' modules of these kinds. In fact, we recently studied the evolutionary combination of loops in domains by generating networks of loops and domains and by tracing their evolution along a timeline of billions of years (Aziz et al., 2016). We uncovered remarkable patterns such as the existence of two functional 'waves' of innovation associated with the 'p-loop' and 'winged helix' general domain structures, the preferential recruitment of ancient loops into new domain structures, and a pervasive network tendency toward hierarchical modularity. Given this difficult 'mosaic' problem that affects the sequences of genes and demands phylogenetic dissection, it is interesting to observe here that the tree of proteomes and the trees reconstructed from central genes provided a same overall phylogenetic insight of four TRUCs.

In summary, we highlight here the quantum leap that exists between classical and giant viruses. Our analyses confirm previous evidence of the existence of a fourth TRUC of life that includes viruses, and highlight its ancestrality and mosaicism. Results suggest that best representations for the evolution of giant viruses and cellular microorganisms are rhizomes, and, beyond, that mosaicism has to be considered at the genome (gene content) level but, more generally, at the gene and sequence level. Giant viruses may be represented as comprised by an evolutionary core inferred from highly conserved protein fold structures and gene sequences of very central and ancient proteins, surrounded by a larger and more dynamic gene complement characterized by genome and gene sequence mosaicisms. Such an abductive path as we use, which is based on phenotypic observations, is propitious to provide novel insight on microbial evolution. The "Fourth TRUC" club should, beyond any doubt, continue to expand in the near future, which may be boosted by using new amoebae as co-culture supports and by implementing high-throughput isolation strategies (Khalil et al., 2016). These giant viruses, as new biological entities, should continue to challenge previous paradigms, and a first step is to describe extensively these parasitic microbes without ribosomes.

AUTHOR CONTRIBUTIONS

DR, PC, PP, GC-A, BLS, and AL designed the experiments. PC, AL, GC-A, and DR wrote the manuscript. PC, AL, VS, AN, and GC-A performed the experiments. All authors analyzed the data and reviewed the manuscript.

FUNDING

This work was supported by a grant from the French State managed by the National Research Agency under the "Investissements d'Avenir (Investments for the Future)" program with the reference ANR-10-IAHU-03 (Méditerranée Infection) and by the région Provence Alpes Côte d'Azur and European funding FEDER PRIMI. Nisrine Chelkha was financially supported through a grant from the Infectiopole Sud Foundation. Research at Illinois was supported by the USDA National Institute of Food and Agriculture, Hatch project 1014249 and a Blue Waters allocation to GC-A.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02668/full#supplementary-material

REFERENCES

Abdelrahman, Y., Ouellette, S. P., Belland, R. J., and Cox, J. V. (2016). Polarized cell division of *Chlamydia trachomatis*. *PLoS Pathog.* 12:e1005822. doi: 10.1371/journal.ppat.1005822

Abergel, C., Legendre, M., and Claverie, J. M. (2015). The rapidly expanding universe of giant viruses: mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev.* 39, 779–796. doi: 10.1093/femsre/fuv037

Abrahao, J., Silva, L., Santos Silva, L., Bou Khalil, J. Y., Rodrigues, R., Arantes, T., et al. (2018). Tupanvirus, a tailed giant virus and distant relative of Mimiviridae, possesses the most complete translational apparatus of the virosphere. *Nat. Commun.* 9:749. doi: 10.1038/s41467-018-03168-1

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Alva, V., Söding, J., and Lupas, A. N. (2015). A vocabulary of ancient peptides at the origin of folded proteins. *eLife* 4:e09410. doi: 10.7554/eLife.09410

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

Andreani, J., Khalil, J. Y. B., Baptiste, E., Hasni, I., Michelle, C., Raoult, D., et al. (2018). Orpheovirus IHUMI-LCC2: a new virus among the giant viruses. *Front. Microbiol.* 8:2643. doi: 10.3389/fmicb.2017.02643

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

Andreeva, A., Howorth, D., Chandonia, J. M., Brenner, S. E., Hubbard, T. J., Chothia, C. et al. (2008). Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* 36, D419–D425. doi: 10.1093/nar/gkm993

Aziz, M. F., Caetano-Anollés, K., and Caetano-Anollés, G. (2016). The early history and emergence of molecular functions and modular scale-free network behavior. *Sci. Rep.* 6:25058. doi: 10.1038/srep25058

Baird, N. L., York, J., and Nunberg, J. H. (2012). Arenavirus infection induces discrete cytosolic structures for RNA replication. *J. Virol.* 86, 11301–11310. doi: 10.1128/JVI.01635-12

Bajrai, L. H., Benamar, S., Azhar, E. I., Robert, C., Levasseur, A., Raoult, D., et al. (2016). Kaumoebavirus, a new virus that clusters with faustoviruses and Asfarviridae. *Viruses* 8:E278. doi: 10.3390/v8110278

Ben, S. I., Adekambi, T., Raoult, D., and Drancourt, M. (2008). RpoB sequence-based identification of Mycobacterium avium complex species. *Microbiology* 154, 3715–3723. doi: 10.1099/mic.0.2008/020164-0

Benamar, S., Reteno, D. G., Bandaly, V., Labas, N., Raoult, D., and La Scola, B. (2016). Faustoviruses: comparative genomics of new Megavirales family members. *Front. Microbiol.* 7:3. doi: 10.3389/fmicb.2016.00003

Bou Khalil, J. Y., Benamar, S., Baudoin, J. P., Croce, O., Blanc-Tailleur, C., Pagnier, I., et al. (2016). Developmental cycle and genome analysis of "*Rubidus massiliensis*," a New *Vermamoeba vermiformis* Pathogen. *Front. Cell. Infect. Microbiol.* 6:31. doi: 10.3389/fcimb.2016.00031

Boyer, M., Madoui, M. A., Gimenez, G., La Scola, B., and Raoult, D. (2010). Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. *PLoS One* 5:e15530. doi: 10.1371/journal.pone.0015530

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21848–21853. doi: 10.1073/pnas.0911354106

Byrne, D., Grzela, R., Lartigue, A., Audic, S., Chenivesse, S., Encinas, S., et al. (2009). The polyadenylation site of Mimivirus transcripts obeys a stringent 'hairpin rule'. *Genome Res.* 19, 1233–1242. doi: 10.1101/gr.091561.109

Caetano-Anollés, G., and Caetano-Anollés, D. (2003). An evolutionarily structured universe of protein architecture. *Genome Res.* 13, 1563–1571. doi: 10.1101/gr.1161903

Caetano-Anollés, G., Kim, K. M., and Caetano-Anollés, D. (2012). The phylogenomic roots of modern biochemistry: origins of proteins, cofactors and protein biosynthesis. *J. Mol. Evol.* 74, 1–34. doi: 10.1007/s00239-011-9480-1

Caetano-Anollés, G., Wang, M., and Caetano-Anollés, D. (2013). Structural phylogenomics retrodicts the origin of the genetic code and uncovers the evolutionary impact of protein flexibility. *PLoS One* 8:e72225. doi: 10.1371/journal.pone.0072225

Caetano-Anollés, G., Wang, M., Caetano-Anollés, D., and Mittenthal, J. E. (2009). The origin, evolution and structure of the protein world. *Biochem. J.* 417, 621–637. doi: 10.1042/BJ20082063

Caetano-Anollés, K., Caetano-Anollés, D., Nasir, A., Kim, K. M., and Caetano-Anollés, G. (2018). Order and polarity in character state transformation models that root the tree of life. *Biochimie* 149, 135–136. doi: 10.1016/j.biochi.2018.04.001

Claverie, J. M., and Abergel, C. (2009). Mimivirus and its virophage. *Annu. Rev. Genet.* 43, 49–66. doi: 10.1146/annurev-genet-102108-134255

Colson, P., de Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D. K., et al. (2013a). "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* 158, 2517–2521. doi: 10.1007/s00705-013-1768-6

Colson, P., Pagnier, I., Yoosuf, N., Fournous, G., La Scola, B., and Raoult, D. (2013b). "Marseilleviridae", a new family of giant viruses infecting amoebae. *Arch. Virol.* 158, 915–920. doi: 10.1007/s00705-012-1537-y

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

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

Condit, R. C., Moussatche, N., and Traktman, P. (2006). In a nutshell: structure and assembly of the vaccinia virion. *Adv. Virus Res.* 66, 31–124. doi: 10.1016/S0065-3527(06)66002-8

Corradi, N., Pombert, J. F., Farinelli, L., Didier, E. S., and Keeling, P. J. (2010). The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nat. Commun.* 1:77. doi: 10.1038/ncomms1082

Dagan, T., and Martin, W. (2006). The tree of one percent. *Genome Biol.* 7:118. doi: 10.1186/gb-2006-7-10-118

Deleuze, G., and Guattari, F. (1976). *Rhizome : Introduction*. Paris: de Minuit.

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

Dou, C., Yu, M., Gu, Y., Wang, J., Yin, K., Nie, C., et al. (2018). Structural and mechanistic analyses reveal a unique Cas4-like protein in the mimivirus virophage resistance element system. *iScience* 3, 1–10. doi: 10.1016/j.isci.2018.04.001

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Fabre, E., Jeudy, S., Santini, S., Legendre, M., Trauchessec, M., Coute, Y., et al. (2017). Noumeavirus replication relies on a transient remote control of the host nucleus. *Nat. Commun.* 8:15087. doi: 10.1038/ncomms15087

Farris, J. S. (1983). "The logical basis of phylogenetic analysis," in *Advances in Cladistics Proceedings of the Second Meeting of the Willi Hennig Society*, Vol. 2, eds N. I. Platnick and V. A. Funk (New York, NY: Columbia University Press), 7–36.

Federici, B. A., Vlak, J. M., and Hamm, J. J. (1990). Comparative study of virion structure, protein composition and genomic DNA of three ascovirus isolates. *J. Gen. Virol.* 71, 1661–1668. doi: 10.1099/0022-1317-71-8-1661

Fischer, M. G., and Suttle, C. A. (2011). A virophage at the origin of large DNA transposons. *Science* 332, 231–234. doi: 10.1126/science.1199412

Forterre, P., and Gaia, M. (2016). Giant viruses and the origin of modern eukaryotes. *Curr. Opin. Microbiol.* 31, 44–49. doi: 10.1016/j.mib.2016.02.001

Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., et al. (2006). Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 103, 425–430. doi: 10.1073/pnas.0510013103

Goncearenco, A., and Berezovsky, I. N. (2015). Protein function from its emergence to diversity in contemporary proteins. *Phys. Biol.* 12:045002. doi: 10.1088/1478-3975/12/4/045002

Gorlas, A., Robert, C., Gimenez, G., Drancourt, M., and Raoult, D. (2012). Complete genome sequence of *Methanomassiliicoccus luminyensis*, the largest genome of a human-associated *Archaea* species. *J. Bacteriol.* 194:4745. doi: 10.1128/JB.00956-12

Gough, J., and Chothia, C. (2002). SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments. *Nucleic Acids Res.* 30, 268–272. doi: 10.1093/nar/30.1.268

Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* 313, 903–919. doi: 10.1006/jmbi.2001.5080

Gould, S. J., and Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B Biol. Sci.* 205, 581–598. doi: 10.1098/rspb.1979.0086

Harish, A., and Caetano-Anollés, G. (2012). Ribosomal history reveals origins of modern protein synthesis. *PLoS One* 7:e32776. doi: 10.1371/journal.pone.0032776

Hedges, S. B., Dudley, J., and Kumar, S. (2006). TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22, 2971–2972. doi: 10.1093/bioinformatics/btl505

Illergard, K., Ardell, D. H., and Elofsson, A. (2009). Structure is three to ten times more conserved than sequence–a study of structural response in protein cores. *Proteins* 77, 499–508. doi: 10.1002/prot.22458

Iyer, L. M., Aravind, L., and Koonin, E. V. (2001). Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.* 75, 11720–11734. doi: 10.1128/JVI.75.23.11720-11734.2001

Iyer, L. M., Balaji, S., Koonin, E. V., and Aravind, L. (2006). Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* 117, 156–184. doi: 10.1016/j.virusres.2006.01.009

Kawai, H., Kanegae, T., Christensen, S., Kiyosue, T., Sato, Y., Imaizumi, T., et al. (2003). Responses of ferns to red light are mediated by an unconventional photoreceptor. *Nature* 421, 287–290. doi: 10.1038/nature01310

Khalil, J. Y., Robert, S., Reteno, D. G., Andreani, J., Raoult, D., and La Scola, B. (2016). High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. *Front. Microbiol.* 7:26. doi: 10.3389/fmicb.2016.00026

Koonin, E. V., and Dolja, V. V. (2014). Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol. Mol. Biol. Rev.* 78, 278–303. doi: 10.1128/MMBR.00049-13

Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2006). The ancient virus world and evolution of cells. *Biol. Direct* 1:29. doi: 10.1186/1745-6150-1-29

Koonin, E. V., and Yutin, N. (2010). Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. *Intervirology* 53, 284–292. doi: 10.1159/000312913

Krupovic, M., and Koonin, E. V. (2017). Cellular origin of the viral capsid-like bacterial microcompartments. *Biol. Direct* 12:25. doi: 10.1186/s13062-017-0197-y

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

La Scola, B., de Lamballerie, X. N., Claverie, J. M., Drancourt, M., and Raoult, D. (2005). "Genus Mimivirus," in *Virus Taxonomy*, eds M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (San Diego: Elsevier Academic Press), 275–276.

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* 455, 100–104. doi: 10.1038/nature07218

Lartillot, N., Brinkmann, H., and Philippe, H. (2007). Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol. Biol.* 7(Suppl. 1):S4. doi: 10.1186/1471-2148-7-S1-S

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

Legendre, M., Arslan, D., Abergel, C., and Claverie, J. M. (2012). Genomics of Megavirus and the elusive fourth domain of Life. *Commun. Integr. Biol.* 5, 102–106. doi: 10.4161/cib.18624

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (2015). In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5327–E5335. doi: 10.1073/pnas.1510795112

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

Levasseur, A., Bekliz, M., Chabriere, E., Pontarotti, P., La Scola, B., and Raoult, D. (2016b). MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature* 531, 249–252. doi: 10.1038/nature17146

Levasseur, A., Merhej, V., Baptiste, E., Sharma, V., Pontarotti, P., and Raoult, D. (2017). The Rhizome of Lokiarchaeota illustrates the mosaicity of archaeal genomes. *Genome Biol. Evol.* 9, 2635–2639. doi: 10.1093/gbe/evx208

Li, L., Stoeckert, C. J. Jr., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503

Long, M., Wang, W., and Zhang, J. (1999). Origin of new genes and source for N-terminal domain of the chimerical gene, *jingwei*, in *Drosophila*. *Gene* 238, 135–141. doi: 10.1016/S0378-1119(99)00229-2

Lundberg, J. G. (1972). Wagner networks and ancestors. *Syst. Zool.* 21, 398–413. doi: 10.1093/sysbio/21.4.398

Lwoff, A. (1957). The concept of virus. *J. Gen. Microbiol.* 17, 239–253. doi: 10.1099/00221287-17-2-239

Meheust, R., Bhattacharya, D., Pathmanathan, J. S., McInerney, J. O., Lopez, P., and Bapteste, E. (2018). Formation of chimeric genes with essential functions at the origin of eukaryotes. *BMC Biol.* 16:30. doi: 10.1186/s12915-018-0500-0

Merhej, V., Notredame, C., Royer-Carenzi, M., Pontarotti, P., and Raoult, D. (2011). The rhizome of life: the sympatric *Rickettsia felis* paradigm demonstrates the random transfer of DNA sequences. *Mol. Biol. Evol.* 28, 3213–3223. doi: 10.1093/molbev/msr239

Moliner, C., Fournier, P. E., and Raoult, D. (2010). Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiol. Rev.* 34, 281–294. doi: 10.1111/j.1574-6976.2010.00209.x

Moreira, D., and Lopez-Garcia, P. (2009). Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* 7, 306–311. doi: 10.1038/nrmicro2108

Moreira, D., and Lopez-Garcia, P. (2015). Evolution of viruses and cells: do we need a fourth domain of life to explain the origin of eukaryotes? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140327. doi: 10.1098/rstb.2014.0327

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., et al. (2006). The 160-kilobase genome of the bacterial endosymbiont Carsonella. *Science* 314:267. doi: 10.1126/science.1134196

Nasir, A., and Caetano-Anollés, G. (2015). A phylogenomic data-driven exploration of viral origins and evolution. *Sci. Adv.* 1:e1500527. doi: 10.1126/sciadv.1500527

Nasir, A., Kim, K. M., and Caetano-Anollés, G. (2012). Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evol. Biol.* 12:156. doi: 10.1186/1471-2148-12-156

Nasir, A., Kim, K. M., and Caetano-Anollés, G. (2017). Phylogenetic tracings of proteome size support the gradual accretion of protein structural domains and the early origin of viruses from primordial cells. *Front. Microbiol.* 8:1178. doi: 10.3389/fmicb.2017.01178

Nasir, A., Sun, F. J., Kim, K. M., and Caetano-Anollés, G. (2015). Untangling the origin of viruses and their impact on cellular evolution. *Ann. N. Y. Acad. Sci.* 1341, 61–74. doi: 10.1111/nyas.12735

Netherton, C. L., and Wileman, T. (2011). Virus factories, double membrane vesicles and viroplasm generated in animal cells. *Curr. Opin. Virol.* 1, 381–387. doi: 10.1016/j.coviro.2011.09.008

Nissimov, J. I., Pagarete, A., Ma, F., Cody, S., Dunigan, D. D., Kimmance, S. A., et al. (2017). Coccolithoviruses: a review of cross-kingdom genomic thievery and metabolic thuggery. *Viruses* 9:E52. doi: 10.3390/v9030052

Ogata, H., La Scola, B., Audic, S., Renesto, P., Blanc, G., Robert, C., et al. (2006). Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet.* 2:e76. doi: 10.1371/journal.pgen.0020076

Pagnier, I., Yutin, N., Croce, O., Makarova, K. S., Wolf, Y. I., Benamar, S., et al. (2015). *Babela massiliensis*, a representative of a widespread bacterial phylum with unusual adaptations to parasitism in amoebae. *Biol. Direct* 10:13. doi: 10.1186/s13062-015-0043-z

Pathmanathan, J. S., Lopez, P., Lapointe, F. J., and Bapteste, E. (2018). Composite Search: a generalized network approach for composite gene families detection. *Mol. Biol. Evol.* 35, 252–255. doi: 10.1093/molbev/msx283

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

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

Raoult, D. (2010). The post-Darwinist rhizome of life. *Lancet* 375, 104–105. doi: 10.1016/S0140-6736(09)61958-9

Raoult, D. (2013). TRUC or the need for a new microbial classification. *Intervirology* 56, 349–353. doi: 10.1159/000354269

Raoult, D. (2014). How the virophage compels the need to readdress the classification of microbes. *Virology* 477, 119–124. doi: 10.1016/j.virol.2014.11.014

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004). The 1.2-megabase genome sequence of Mimivirus. *Science* 306, 1344–1350. doi: 10.1126/science.1101485

Raoult, D., and Boyer, M. (2010). Amoebae as genitors and reservoirs of giant viruses. *Intervirology* 53, 321–329. doi: 10.1159/000312917

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

Raoult, D., La Scola, B., and Birtles, R. (2007). The discovery and characterization of Mimivirus, the largest known virus and putative pneumonia agent. *Clin. Infect. Dis.* 45, 95–102. doi: 10.1086/518608

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

Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., et al. (2017). Giant viruses with an expanded complement of translation system components. *Science* 356, 82–85. doi: 10.1126/science.aal4657

Schvarcz, C. R., and Steward, G. F. (2018). A giant virus infecting green algae encodes key fermentation genes. *Virology* 518, 423–433. doi: 10.1016/j.virol.2018.03.010

Seligmann, H., and Raoult, D. (2016). Unifying view of stem-loop hairpin RNA as origin of current and ancient parasitic and non-parasitic RNAs, including in giant viruses. *Curr. Opin. Microbiol.* 31, 1–8. doi: 10.1016/j.mib.2015.11.004

Sharma, V., Colson, P., Chabrol, O., Pontarotti, P., and Raoult, D. (2015a). *Pithovirus sibericum*, a new bona fide member of the "Fourth TRUC" club. *Front. Microbiol.* 6:722. doi: 10.3389/fmicb.2015.00722

Sharma, V., Colson, P., Chabrol, O., Scheid, P., Pontarotti, P., and Raoult, D. (2015b). Welcome to pandoraviruses at the 'Fourth TRUC' club. *Front. Microbiol.* 6:423. doi: 10.3389/fmicb.2015.00423

Sharma, V., Colson, P., Giorgi, R., Pontarotti, P., and Raoult, D. (2014). DNA-dependent RNA polymerase detects hidden giant viruses in published databanks. *Genome Biol. Evol.* 6, 1603–1610. doi: 10.1093/gbe/evu128

Sharma, V., Colson, P., Pontarotti, P., and Raoult, D. (2016). Mimivirus inaugurated in the 21st century the beginning of a reclassification of viruses. *Curr. Opin. Microbiol.* 31, 16–24. doi: 10.1016/j.mib.2015.12.010

Swofford, D. L. (2002). *PAUP*, Version 4.0b10*. Sunderland, MA: Sinauer.

Tamames, J., Gil, R., Latorre, A., Pereto, J., Silva, F. J., and Moya, A. (2007). The frontier between cell and organelle: genome analysis of Candidatus *Carsonella ruddii*. *BMC Evol. Biol.* 7:181. doi: 10.1186/1471-2148-7-181

Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C., et al. (2008). Atomic-level models of the bacterial carboxysome shell. *Science* 319, 1083–1086. doi: 10.1126/science.1151458

Terhune, S. S., Schroer, J., and Shenk, T. (2004). RNAs are packaged into human cytomegalovirus virions in proportion to their intracellular concentration. *J. Virol.* 78, 10390–10398. doi: 10.1128/JVI.78.19.10390-10398.2004

Wang, M., Jiang, Y. Y., Kim, K. M., Qu, G., Ji, H. F., Mittenthal, J. E., et al. (2011). A universal molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. *Mol. Biol. Evol.* 28, 567–582. doi: 10.1093/molbev/msq232

Williams, T. A., Embley, T. M., and Heinz, E. (2011). Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses. *PLoS One* 6:e21080. doi: 10.1371/journal.pone.0021080

Wilson, W. H., Van Etten, J. L., and Allen, M. J. (2009). The Phycodnaviridae: the story of how tiny giants rule the world. *Curr. Top. Microbiol. Immunol.* 328, 1–42. doi: 10.1007/978-3-540-68618-7_1

Wu, D., Wu, M., Halpern, A., Rusch, D. B., Yooseph, S., Frazier, M., et al. (2011). Stalking the fourth domain in metagenomic data: searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees. *PLoS One* 6:e18011. doi: 10.1371/journal.pone.0018011

Yutin, N., and Koonin, E. V. (2012). Hidden evolutionary complexity of nucleo-cytoplasmic large DNA viruses of eukaryotes. *Virol. J.* 9:161. doi: 10.1186/1743-422X-9-161

Yutin, N., and Koonin, E. V. (2013). Pandoraviruses are highly derived phycodnaviruses. *Biol. Direct* 8, 25–28. doi: 10.1186/1745-6150-8-25

Yutin, N., Wolf, Y. I., and Koonin, E. V. (2014). Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* 46, 38–52. doi: 10.1016/j.virol.2014.06.032

Yutin, N., Wolf, Y. I., Raoult, D., and Koonin, E. V. (2009). Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol. J.* 17:223. doi: 10.1186/1743-422X-6-223

Zhang, J. Z., Popov, V. L., Gao, S., Walker, D. H., and Yu, X. J. (2007). The developmental cycle of *Ehrlichia chaffeensis* in vertebrate cells. *Cell. Microbiol.* 9, 610–618. doi: 10.1111/j.1462-5822.2006.00812.x

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 September 2018
doi: 10.3389/fmicb.2018.02098

A Phylogenomic Study of *Acanthamoeba polyphaga* Draft Genome Sequences Suggests Genetic Exchanges With Giant Viruses

Nisrine Chelkha, Anthony Levasseur, Pierre Pontarotti, Didier Raoult, Bernard La Scola and Philippe Colson*

Institut de Recherche pour le Développement, Assistance Publique – Hôpitaux de Marseille, Microbes, Evolution, Phylogeny and Infection, and Institut Hospitalo-Universitaire – Méditerranée Infection, Aix-Marseille Université, Marseille, France

Acanthamoeba are ubiquitous phagocytes predominant in soil and water which can ingest many microbes. Giant viruses of amoebae are listed among the *Acanthamoeba*-resisting microorganisms. Their sympatric lifestyle within amoebae is suspected to promote lateral nucleotide sequence transfers. Some *Acanthamoeba* species have shown differences in their susceptibility to giant viruses. Until recently, only the genome of a single *Acanthamoeba castellanii* Neff was available. We analyzed the draft genome sequences of *Acanthamoeba polyphaga* through several approaches, including comparative genomics, phylogeny, and sequence networks, with the aim of detecting putative nucleotide sequence exchanges with giant viruses. We identified a putative sequence trafficking between this *Acanthamoeba* species and giant viruses, with 366 genes best matching with viral genes. Among viruses, Pandoraviruses provided the greatest number of best hits with 117 (32%) for *A. polyphaga*. Then, genes from mimiviruses, Mollivirus sibericum, marseilleviruses, and Pithovirus sibericum were best hits in 67 (18%), 35 (9%), 24 (7%), and 2 (0.5%) cases, respectively. Phylogenetic reconstructions showed in a few cases that the most parsimonious evolutionary scenarios were a transfer of gene sequences from giant viruses to *A. polyphaga.* Nevertheless, in most cases, phylogenies were inconclusive regarding the sense of the sequence flow. The number and nature of putative nucleotide sequence transfers between *A. polyphaga,* and *A. castellanii* ATCC 50370 on the one hand, and pandoraviruses, mimiviruses and marseilleviruses on the other hand were analyzed. The results showed a lower number of differences within the same giant viral family compared to between different giant virus families. The evolution of 10 scaffolds that were identified among the 14 *Acanthamoeba* sp. draft genome sequences and that harbored $\geq$ 3 genes best matching with viruses showed a conservation of these scaffolds and their 46 viral genes in *A. polyphaga*, *A. castellanii* ATCC 50370 and *A. pearcei*. In contrast, the number of conserved genes decreased for other *Acanthamoeba* species, and none of these 46 genes were present in three of them. Overall, this work opens up several potential avenues for future studies on the interactions between *Acanthamoeba* species and giant viruses.

Keywords: *Acanthamoeba polyphaga*, *Acanthamoeba*, giant viruses, mimivirus, draft genome sequences, horizontal gene transfer, nucleotide sequence transfer

OPEN ACCESS

Edited by:
Akio Adachi,
Kansai Medical University, Japan

Reviewed by:
Kiran Kondabagil,
Indian Institute of Technology Bombay, India
Carsten Balczun,
Bundeswehrkrankenhaus Koblenz, Germany
Erna Geessien Kroon,
Universidade Federal de Minas Gerais (UFMG), Brazil

***Correspondence:**
Philippe Colson
philippe.colson@univ-amu.fr

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

Received: *14 June 2018*
Accepted: *16 August 2018*
Published: *06 September 2018*

Citation:
Chelkha N, Levasseur A, Pontarotti P, Raoult D, La Scola B and Colson P (2018) A Phylogenomic Study of Acanthamoeba polyphaga Draft Genome Sequences Suggests Genetic Exchanges With Giant Viruses. Front. Microbiol. 9:2098. doi: 10.3389/fmicb.2018.02098

INTRODUCTION

Acanthamoeba spp. (*Eukaryota*, *Amoebozoa*, *Acanthamoebidae*) are among the most predominant protozoa in soil and water (Rodríguez-Zaragoza, 1994). These amoebae are found in natural or artificial habitats, mostly humid ones, such as the marine environment, sediments, salt lakes, cooling towers, stagnant water, treatment plant sewage, drinking water, or soil (Rodríguez-Zaragoza, 1994). Their ubiquity in water and soil promotes contacts with animals including humans (Sherr and Sherr, 2002; Marciano-cabral and Cabral, 2003). Moreover, *Acanthamoeba* spp. are phagocytic protists that can ingest all particles with a size > 0.5 μm, which includes notably bacteria and fungi (Raoult and Boyer, 2010). Whereas most of these microorganisms are degraded post-internalization, some are able to survive and multiply (Raoult and Boyer, 2010). They are known as amoeba-resisting microorganisms (ARMs) (Greub and Raoult, 2004). Examples of ARMs include human pathogens such as *Legionella pneumophila* or *Mycobacterium* sp., which also resist degradation by macrophages (Greub and Raoult, 2004; Salah et al., 2009).

Due to their giant size, giant viruses of amoebae are ARMs that multiply inside *Acanthamoeba polyphaga* and *A. castellanii* (Colson et al., 2017b). These amoebae have been used in laboratory settings for giant virus isolation during the last 14 years since the discovery of their first representative, Mimivirus (La Scola et al., 2003; Raoult et al., 2004; Colson et al., 2017a). Mimivirus led to the creation of the family *Mimiviridae* and to the discovery of many other giant viruses of amoebae, as well as virophages that replicate in mimivirus factories, and transpovirons integrated in mimivirus genomes (Colson et al., 2017a). Until now, three new families of amoeba-infecting viruses have been described, which include *Mimiviridae*, *Marseilleviridae*, and *Lavidaviridae*, as well as seven new putative lineages consisting in pandoraviruses, pithoviruses, faustoviruses, Mollivirus sibericum, Kaumoebavirus, cedratviruses and Pacmanvirus (Colson et al., 2017b). These giant viruses are commonly found in environmental water and soil and have a broad geographical distribution. They differ from classical viruses and have a complexity similar to that of other microbes. Emblematically, the virions are visible with an optical microscope and can reach 1.5 μm in size, and their genomes harbor between 444 (for a marseillevirus) and 2,544 (for a pandoravirus) genes (Raoult et al., 2004; Colson et al., 2017a).

The hosting by *Acanthamoeba* spp. of several ARMs living sympatrically confers to these microorganisms increased opportunities to exchange sequences between each other and with the amoebal host (Greub and Raoult, 2004; Raoult and Boyer, 2010; Bertelli and Greub, 2012). This is suspected to promote the broad mosaicism and large size of the genomes of giant viruses of amoebae that are in close vicinity with other viruses and microorganisms inside *Acanthamoeba*. It has also been observed that bacteria and viruses living in sympatry in *Acanthamoeba* harbored larger genomes than their closest relatives with an allopatric lifestyle (Raoult and Boyer, 2010). Such sequence exchanges may allow accumulating a substantial gene armory to multiply and compete with other amoeba-resisting microorganisms (Boyer et al., 2011). Consistently, culturing Mimivirus in allopatric conditions on ARM-free *Acanthamoeba* led to a 16% reduction of the viral genome after 100 passages (Boyer et al., 2011).

According to recent findings, the species *A. castellanii* and *A. polyphaga*, which were those used to isolate giant viruses of amoebae, have different levels of tolerance to viruses from different or even similar lineages (Dornas et al., 2015). We observed that some virus lineages relied on specific *Acanthamoeba* species for their replication. For example, pandoraviruses and pithoviruses were only isolated on *A. castellanii* and faustoviruses were only isolated on *Vermamoeba vermiformis* (Khalil et al., 2016; Reteno et al., 2015). Moreover, different mimivirus isolates were obtained from the same sample with different *Acanthamoeba* species (Dornas et al., 2015). These data raise some questions about the relationship between *Acanthamoeba* species and giant viruses of amoebae.

Giant viruses of amoebae have raised a radically new issue regarding their genomic content. Since the description of the Mimivirus, the question of the origin of their genes has arisen. From the onset, genes corresponding to nucleotide sequences putatively transferred from amoebae to viruses were identified, and it was proposed that giant viruses were essentially bags of exogenous genes (Filée et al., 2007; Moreira and Brochier-Armanet, 2008). This assertion is only partially true, since several genes were suspected to be shared by several giant viruses, and the proportion of ORFan genes, for which no source is identified, remains very high in giant viruses (Colson et al., 2017b). Nucleotide sequences from bacterial viruses (bacteriophages), viruses of viruses (virophages), archaeal viruses, or eukaryotic viruses can possibly be transferred inside their host genomes. Regarding amoebae, at present, nucleotide sequence transfers are mostly studied by comparing their genomes to those of giant viruses. *A. castellanii* was the only *Acanthamoeba* species for which we had draft genome sequences (from 2010 to 2015), which was completed using transcriptomic data (Clarke et al., 2013). The study of this genome led to infer sequence exchanges with other eukaryotes and with archaea, bacteria, and viruses of the proposed order Megavirales, which comprises the formerly described Nucleocytoplasmic large DNA viruses (NCLDVs) and giant viruses of amoebae (Yutin et al., 2009; Colson et al., 2013). In this work, we sought to detect and characterize putative nucleotide sequence transfers involving giant virus genomes and the draft genome sequences of a second *Acanthamoeba* species, *A. polyphaga*, which has been one of the most used to isolate giant viruses of amoebae.

MATERIALS AND METHODS

Gene Content of Giant Viruses

The study of the gene trafficking between *A. polyphaga* and giant viruses was carried out by using genes of all annotated genomes described for these viruses at the time of our analysis (including mimiviriruses, marseilleviruses, pandoraviruses, Pithovirus sibericum, Mollivirus sibericum, faustoviruses, phycodnaviruses, ascoviruses, iridoviruses, asfarviruses and poxviruses) (**Supplementary Table S1**).

Draft Genome Sequences of *Acanthamoeba polyphaga* and Other *Acanthamoeba* Species

The draft genome of *A. polyphaga* ATCC 30872 is publicly available on the NCBI website[1] (accession: PRJEB7687). It is part of the project «Phylogenomics of *Acanthamoeba* species» (Institute of Integrative Biology, University of Liverpool), along with the draft genomes of 13 other *Acanthamoeba* species. The *A. polyphaga* ATCC 30872 draft genome contains 224,482 scaffolds with a total length of 120.6 megabases (Mb). The 14,974 genes identified in the genome of *A. castellanii* Neff with the support of transcriptomic data (Clarke et al., 2013) were used for comparative genomic analyses.

Optimization of Assembly of the *Acanthamoeba* Draft Genome Sequences

We used the CLC Genomics Workbench software (version 7.5)[2] to reassemble the draft genome sequences of *A. polyphaga* and some other *Acanthamoeba* species, including another *A. castellanii* strain (ATCC 50370). The default kmer size of 50 was used.

Gene Prediction, Functional Annotation, and Analysis of Taxonomical Distribution

Gene prediction for the draft genome sequences of *A. polyphaga* and other *Acanthamoeba* species was performed using the Prodigal program, as we searched for sequences best matching those of giant viruses. This tool identifies ribosome binding sites to localize translation initiation positions and localizes precisely the 3′ end of each gene (Hyatt et al., 2010). The hits identified as those best matching with giant virus genes were then checked and compared with the 47,246 genes predicted for *A. polyphaga* using GeneMarkES, a program developed specifically for eukaryotes (Lomsadze et al., 2005). For the functional annotation, sequences homologous with ORFs predicted from non-redundant scaffolds were searched for in the NCBI GenBank protein sequence database (nr) using the BLASTp program (Altschul et al., 1990). In addition, ORFs were identified through BLASTp searches (with an *e*-value threshold of 0.1) in the database of Clusters of Orthologous Groups of proteins (COGs) of the NCBI (Tatusov, 2000). Taxonomical origins were determined using MEGAN6 (Huson et al., 2016).

Comparative Genomic Analyses

Predicted protein sequences of *A. polyphaga* were compared with those from giant viruses and those predicted from the draft genome sequences of the 13 other species of *Acanthamoeba* (**Supplementary Table S2**). The orthologous genes ($\geq$100 amino acids) of *A. polyphaga* and the other *Acanthamoeba* were identified using the Proteinortho program (Lechner et al., 2011). Phylogenetic reconstructions were then performed using two sets of aligned sequences. First, 16 draft genome sequences of *Acanthamoeba* strains classified in 14 species including two *A. polyphaga* strains and two *A. castellanii* strains were aligned by using the progressive Mauve program (Darling et al., 2010). Second, the 18S ribosomal DNA sequences from 33 different *Acanthamoeba* strains were retrieved, including those of the 16 *Acanthamoeba* strains classified in 14 *Acanthamoeba* species. The 18S ribosomal DNA sequences were obtained by sequencing the complete 18S ribosomal DNA from strains available in our laboratory, and for non-available strains, the sequences were retrieved from the NCBI GenBank database or directly from the *Acanthamoeba* draft genome sequences. Nucleotide sequences alignments were performed with the MUSCLE program (Edgar, 2004). Phylogenetic trees were constructed using FastTree (Price et al., 2010).

We investigated more specifically any possible occurrences of horizontal gene transfers (HGT), i.e., the gene trafficking between this amoeba and giant viruses. *A. polyphaga* proteins which had as significant hit a giant virus sequence were used as queries to search into the NCBI GenBank non-redundant protein sequence database (nr). Phylogenetic analyses were performed to confirm suspicion of HGT for the genes showing the highest level of sequence similarity with a viral homolog. Amino acid sequences alignments were performed with the MUSCLE program. Phylogenetic trees were constructed using FastTree. Ancestral major capsid protein (MCP) sequences were predicted using the MEGA6 program[3]. Additionally, similarity searches were performed by using the tBLASTn program for all ORFs of giant viruses, virophages and transpovirons against the draft genome sequences of *A. polyphaga*. The following criteria were used: a percentage of amino acid identity $\geq$ 30%; an *e*-value $\leq$ 1e-2; and a percentage of coverage of aligned sequences $\geq$ 30%. Finally, results from similarity searches were formatted to create networks of gene trafficking using the Cytoscape tool (Smoot et al., 2011). This software was also used to generate a network between protein sequences from two giant viruses of amoebae, Pandoravirus dulcis and Pandoravirus salinus, and from the draft genome sequences of *A. polyphaga* and *A. castellanii* ATCC 50370. Finally, a 'rhizome' of genes was determined for a few *A. polyphaga* genes whose best hit was a giant virus. This information was obtained by performing BLASTp searches with fragments of this gene obtained by fenestrating its amino acid sequence with a window of 40 amino acids and a sliding step of 20 amino acids. The representation of the mosaicism of these genes was built using the Circos tool[4].

We identified 115 ORFs (i) best matching with a giant virus gene, (ii) larger than 100 amino acids, and (iii) present in scaffolds that harbor a majority of 'non-viral' ORFs. For these 115 ORFs, we performed BLASTp searches against the GenBank protein sequence database nr and a tBLASTn search against all 14 *Acanthamoeba* draft genome sequences. Then, we merged the results of these two BLAST searches by creating a database with significant sequence hits (*e*-value < 1e-4, length of query and subject sequence alignments > 100 amino acids), and performed

[1] http://www.ncbi.nlm.nih.gov/bioproject/

[2] https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/

[3] http://www.megasoftware.net/

[4] http://circos.ca/

another BLASTp search for the 115 ORFs against this database. We thereafter examined if best hits were exclusively or in majority sequences from giant viruses or *Acanthamoeba* spp. or both and performed a phylogenetic reconstruction to determine the sense of the nucleotide sequence transfers between these organisms.

Synteny Analysis in the Draft Genome Sequences of the 14 *Acanthamoeba* Species of *A. polyphaga* Genes for Which the Best Match Is a Giant Virus Gene

Synteny preservation of genes from *A. polyphaga* for which the best match is a giant virus gene was evaluated in the draft genome sequences available for the 13 other *Acanthamoeba* species. For this purpose, we selected 10 *A. polyphaga* scaffolds that carried at least 3 such genes, then searched for scaffolds harboring similar genomic sequences in other *Acanthamoeba* species and strains. Finally, we determined whether *A. polyphaga* genes for which the best match is a giant virus gene were present in the genome scaffolds from other species and strains, and whether these genes were in synteny in the different *Acanthamoeba* draft genome sequences.

RESULTS

Improvement of the Assembly of the *Acanthamoeba* Draft Genome Sequences

The estimated size of the *A. polyphaga* draft genome was 120.6 Mb, compared to 115.3 Mb for the draft genome sequences of *A. castellanii* ATCC 50370, and 41 Mb for the *A. castellanii* Neff genome (Clarke et al., 2013). Using CLC Genomics Workbench, the number of contigs for the *A. polyphaga* draft genome sequences was reduced from 224,482 to 56,709. Contig number reduction was in the same order of magnitude for *A. castellanii* ATCC 50370 (from 221,748 to 56,469) (**Table 1**). We obtained a statistically significant reduction in the average number (±standard deviation) of contigs for the 14 draft genome sequences of *Acanthamoeba* ($p < 1e\text{-}3$, ANOVA test) (**Supplementary Figure S1**).

Acanthamoeba polyphaga Gene Content and Comparison With the Gene Content of Other *Acanthamoeba* Species

Gene prediction performed for the *A. polyphaga* draft genome sequences detected a substantial number of ORFs (equal to 310,496) shorter than 50 amino acids (but greater than 35 amino acids). The number of predicted ORFs with a size comprised between 50 and 100 amino acids or larger than 100 amino acids was also considerable, being of 223,728 and 97,092, respectively (**Supplementary Table S2**). Comparison with the *A. castellanii* Neff gene set showed that 97.2% of its genes were detected in the *A. polyphaga* draft genome sequences. The same proportion (97.2%) of *A. castellanii* Neff genes was detected in the draft genome sequences of *A. castellanii* ATCC 50370. These results indicate that a large majority of *A. castellanii* Neff transcribed genes are present in these *Acanthamoeba* sp. strains (**Supplementary Table S3**). The 421 *A. castellanii* Neff genes that were not detected in the *A. polyphaga* draft genome sequences included mostly genes encoding hypothetical proteins [383 genes (91%)]. Other genes included proteins encoding the six MCP described in the *A. castellanii* Neff genome (Maumus and Blanc, 2016), three NAD-dependent epimerase/dehydratase family proteins, two HEAT repeat domains containing proteins, two Rho termination factor domain containing proteins, a chitin synthase, and a putative autophagy protein (**Supplementary Table S4**). In addition, 48,583 (78.6%) of the 61,786 orthologous groups of genes found in the *A. polyphaga* draft genome sequences were also detected in the *A. castellanii* ATCC 50370 draft genome sequences (**Supplementary Figure S2**). The number of non-ORFan genes and ORFan genes larger than 100 amino acids in the *A. polyphaga* draft genome sequences was equal to 58,185 (70.4%) and 24,484 (29.6%), respectively (**Table 2**). Similar proportions were found for the draft genome sequences of *A. castellanii* ATCC 50370, with 56,920 ORFan genes (69.2%) and 25,390 non-ORFan genes (30.8%).

Phylogenetic reconstruction based on 18S ribosomal genes (**Supplementary Figure S3**) and the tree based on similarities between the 14 draft genome sequences and the genome of *A. castellanii* Neff (**Supplementary Figure S4**) showed that the *A. castellanii* ATCC 50370 and Neff are not clustered together. In contrast, the *A. castellanii* ATCC 50370 is clustered with *A. polyphaga* and *A. pearci*, *A. pearci* being the isolate closest to *A. polyphaga*. These findings suggest that the genomes of *A. castellanii* Neff and those of *A. castellanii* ATCC 50370 belong to different species. We further checked for similarities between 18S ribosomal DNA sequences from the draft genome sequences of *A. polyphaga* analyzed here and the sequence AY026244 from *A. polyphaga* ATCC 30872. We observed that 18S ribosomal DNA sequences from both strains were not clustered together, which questions the accuracy of the identification of one or both genomes.

Taxonomical Distribution of *Acanthamoeba polyphaga* Genes and Possible Gene Trafficking Between *Acanthamoeba* spp. and Giant Viruses

The taxonomical distribution of the best BLAST hits obtained for the *A. polyphaga* proteins indicated that 43% belong to *Amoebozoa* (98% of them belonging to *A. castellanii* Neff), 3% belong to eukaryotes other than *Amoebozoa*, 3% belong to bacteria, 3% to archaea, and 51% were identified to be ORFans. A total of 366 genes (0.07%) had viral genes as best match, which suggests an important gene trafficking between amoebae and their infecting viruses (**Figures 1**, **2A**). A total of 41 (11%) of the 366 viral genes in *A. polyphaga* were found to match with genes transcribed in *A. castellanii* Neff. The functions of these 366 *A. polyphaga* genes are mostly related to replication, recombination and repair [COG category L (18%)]; then signal transduction [T (16%)]; general function [R (15%)]; post-translational modification [O (12%)]; transcription

TABLE 1 | Assembly statistics for the draft genome sequences of the two species *A. polyphaga* and *A. castellanii* using the CLC software.

	A. polyphaga		*A. castellanii*	
	Initial draft genome	**New assembly**	**Initial draft genome**	**New assembly**
Genome size (Mb)	120.6	123.8	115.3	121.2
Number of contigs	224,482	56,709	221,748	56,469
Mean length (nucleotides)	519	2147	536	2184
N50	1,362	3,982	1,454	4,103
Proportion of contigs < 1,000 bp (%)	86.9	44.1	86.8	44.2
Proportion of contigs > 10,000 bp (%)	0.19	2.56	0.26	2.80
Proportion of contigs < 100,000 bp (%)	0.0009	0	0.0004	0.0018

bp, base pairs; N50, 50% of the genome is in contigs larger than this size.

TABLE 2 | Gene annotation for species *A. polyphaga* and *A. castellanii*.

Scaffolds/ORFs	***Acanthamoeba polyphaga***	***Acanthamoeba castellanii***
Scaffolds (draft genome)	224,482	221,748
ORFs annotated (> 100 aa)	58,185	56,920
ORFans (> 100 aa)	24,484	25,390

aa, amino acids. The number of annotated ORFs as well as ORFans, considering only the sequences larger than 100 amino acids.

[K (9%)]; and molecule transport and metabolism [E, F, G, H and P (17%)] (**Table 3**). Most of these 366 genes belong to giant viruses of the proposed order Megavirales. Pandoraviruses were the giant viruses that provided the greatest number of best hits with 117 (32%) for *A. polyphaga*. Then, genes from mimiviruses, Mollivirus sibericum, marseilleviruses, and Pithovirus sibericum, were best hits in 67 (18%), 37 (10%), 24 (7%), and 2 (0.5%) cases, respectively. Among other viral genes, there were 34 genes from phycodnaviruses. Among other viral genes, there were 29 genes from phycodnaviruses. A similar number of viral genes (356) was identified in the draft genome sequences of *A. polyphaga* than in those of *A. castellanii* ATCC 50370 (**Supplementary Table S5**). The 47,246 genes predicted for *A. polyphaga* by using the GeneMarkES program were compared to the 366 genes best matching with virus genes, and this comparison showed that coverage inferior to 50% was only observed for 25 genes (6.8%), while 67 genes did not match with any of the available 47,246 genes. A same analysis was specifically carried out using only available ORFomes from all giant viruses isolated on *Acanthamoeba* spp., allowing the identification of 1,797 genes in the *A. polyphaga* draft genome sequences (1.9% of the 97,092 ORFs larger than 100 amino acids) with a giant viral homolog. In accordance with our previous findings, pandoravirus genes were the most abundant of these giant viral homologous genes, before mimiviruses, marseilleviruses, other giant viruses of amoebae, and phycodnaviruses. A substantial level of gene trafficking between *A. polyphaga* and giant viruses was illustrated by a sequence network that made it possible to observe the number and the nature of genes involved in putative nucleotide sequence transfers (**Figure 2B**).

A total of 262 *A. polyphaga* ORFs had as best hit a coding sequence of a giant virus of amoeba, among which 134 are larger than 100 amino acids and 115 are present in scaffolds harboring a majority of 'non-viral' ORFs. For these 115 ORFs, results from BLAST searches against the non-redundant GenBank protein sequence database and all 14 *Acanthamoeba* draft genome sequences showed different patterns of best hits, including sequence sets comprising a majority of sequences from giant viruses or from *Acanthamoeba* spp. Subsequent phylogenetic reconstructions enabled us to infer that the most parsimonious evolutionary scenario was, in at least three cases, a gene sequence transfer from giant viruses of amoebae to *Acanthamoeba* spp., although alternative scenarios could not be ruled out (**Figures 3A–C**). Also, in at least three cases, gene sequence transfer was supposed to have occurred in the opposite way, from *Acanthamoeba* spp. to giant viruses of amoebae (**Figures 3D–F**). Nevertheless, phylogenies were most often inconclusive regarding the putative sense of the gene flow (**Supplementary Figure S5**). We analyzed further two cases for which the putative sense of the gene sequence transfer was from giant viruses to amoebae and two cases for which the putative sense of the gene sequence transfer was from amoebae to giant viruses. We searched for the most similar sequences for short fragments of these genes. We found that the best hits for these fragments were organisms that belonged to different cellular domains (*Eukarya*, *Bacteria*, or *Archaea*), or were of putative viral origin (**Figure 4**). This gene sequence mosaicism was observed in all four cases, although with level differences. This indicates that sequence mosaicism may also occur within genes, and may challenge the interpretation of gene-based phylogeny.

As previous culture isolation experiments suggested different levels of permissivity to giant viruses according to the *Acanthamoeba* species (Dornas et al., 2015), sequences detected in *A. polyphaga* and *A. castellanii* ATCC 50370 that were homologous with giant viral genes were compared (**Supplementary Table S6** and **Supplementary Figure S6**). This showed that the numbers and sets of viral genes that are homologs of genes present in these two *Acanthamoeba* species differ between giant viral families. In contrast, for a given viral family, a majority of genes found as best matches of genes from these two *Acanthamoeba* species

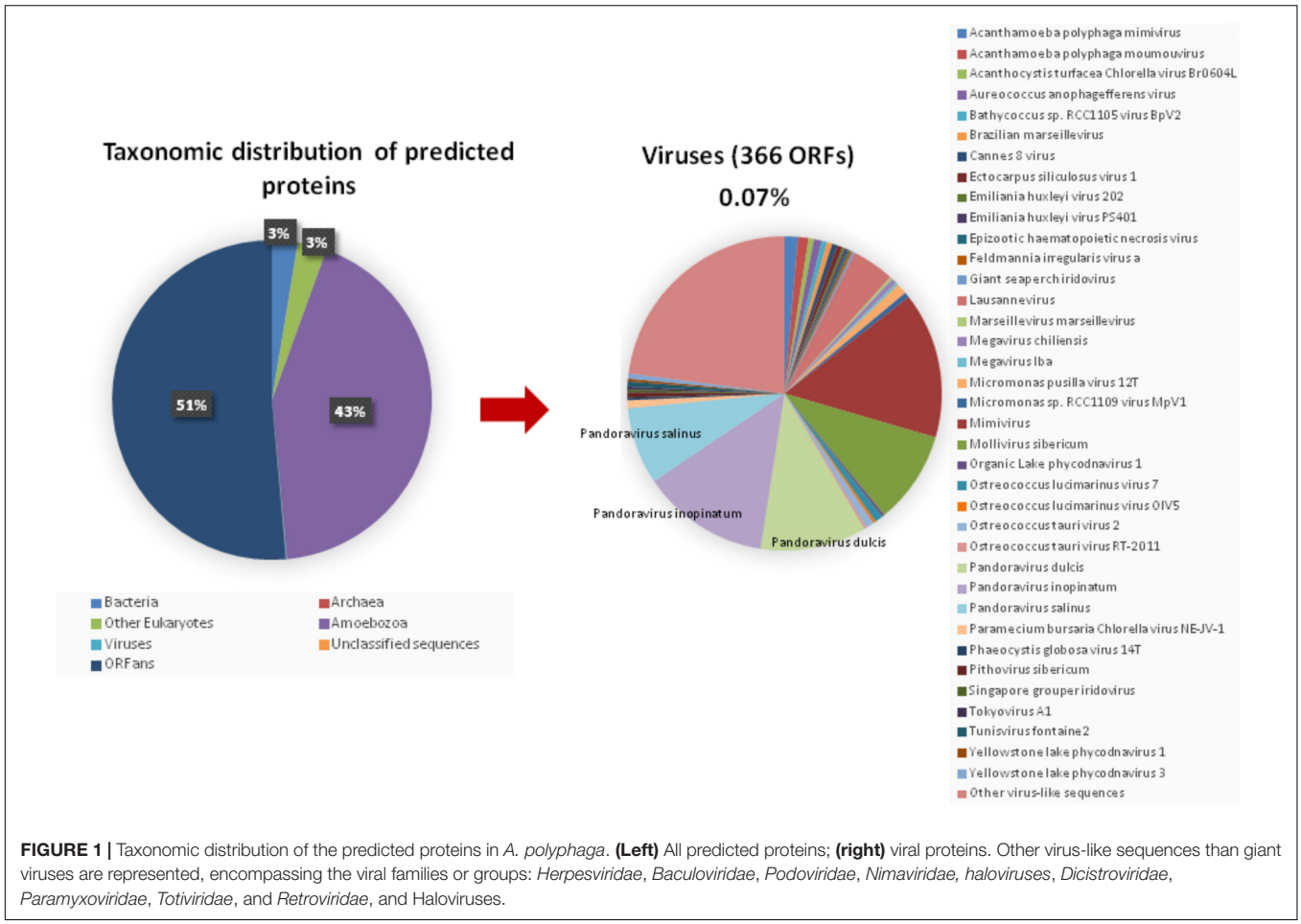


FIGURE 1 | Taxonomic distribution of the predicted proteins in *A. polyphaga*. **(Left)** All predicted proteins; **(right)** viral proteins. Other virus-like sequences than giant viruses are represented, encompassing the viral families or groups: *Herpesviridae*, *Baculoviridae*, *Podoviridae*, *Nimaviridae, haloviruses*, *Dicistroviridae*, *Paramyxoviridae*, *Totiviridae*, and *Retroviridae*, and Haloviruses.

were conserved in different viruses. Nonetheless, in some viruses, we identified genes that are homologous with only one of the two *Acanthamoeba* species. For example, homologs of some mimivirus genes were specifically found in *A. polyphaga* (**Supplementary Table S6**). Furthermore, analysis of the presence and conservation of these genes in the draft genome sequences of the other *Acanthamoeba* species showed that some were present in a majority of these genomes, being only absent in those of two or three other *Acanthamoeba* species. For instance, Pandoravirus salinus gene Ps_2278 was only absent in *A. castellanii* ATCC 50370 and *A. healyi*, whereas Pandoravirus dulcis gene Pd_13-16 was absent in *A. polyphaga*, *A. mauritaniensis,* and *A. pearcei*. In contrast, no giant virus homologous genes were found in the draft genome sequences of a substantial number of *Acanthamoeba* species, including pandoravirus genes Ps_1170 and Pd_589. In addition, the same gene in the draft genome sequences of *A. castellanii* ATCC 50370 was homologous with both Pandoravirus salinus gene Ps_2319 and Pandoravirus dulcis gene Pd_1426 (**Supplementary Table S7**).

Acanthamoeba polyphaga is one of the *Acanthamoeba* species for which no MCP homologs were detected. In contrast, eight MCP homologs were found in the draft genome sequences of other *Acanthamoeba* species including *A. lenticulata*, *A. lugdunensis*, *A. quina*, *A. healyi,* and *A. mauritaniensis* (**Supplementary Figure S7** and **Supplementary Table S8**), as previously described (Maumus and Blanc, 2016). In addition, *A. castellanii* Neff was previously found to harbor six genes encoding MCPs of giant viruses (Maumus and Blanc, 2016). A phylogenetic analysis showed that three of these sequences, including MCP homolog 1, MCP homolog 2, and iridovirus MCP homolog 2, were clustered. Moreover, the iridovirus homolog 1 sequence is clustered with a sequence from Mollivirus sibericum, the nearest neighbor of the pandoraviruses for which no gene encoding a capsid protein has been identified (Legendre et al., 2015) (**Supplementary Figure S8a**). A BLAST search was performed against the NCBI non-redundant protein sequence database. The query used was the ancestral sequence inferred for the MCP protein detected in the draft genome sequences of the different *Acanthamoeba* species and the MCP of Mollivirus sibericum. This search retrieved MCP-encoding sequences from phycodnaviruses, which are DNA viruses that infect algae and are classified with giant viruses of *Acanthamoeba* in the proposed order Megavirales (**Supplementary Figure 8b**). The G + C content was homogenous along the *Acanthamoeba* scaffolds carrying these virus-related genes.

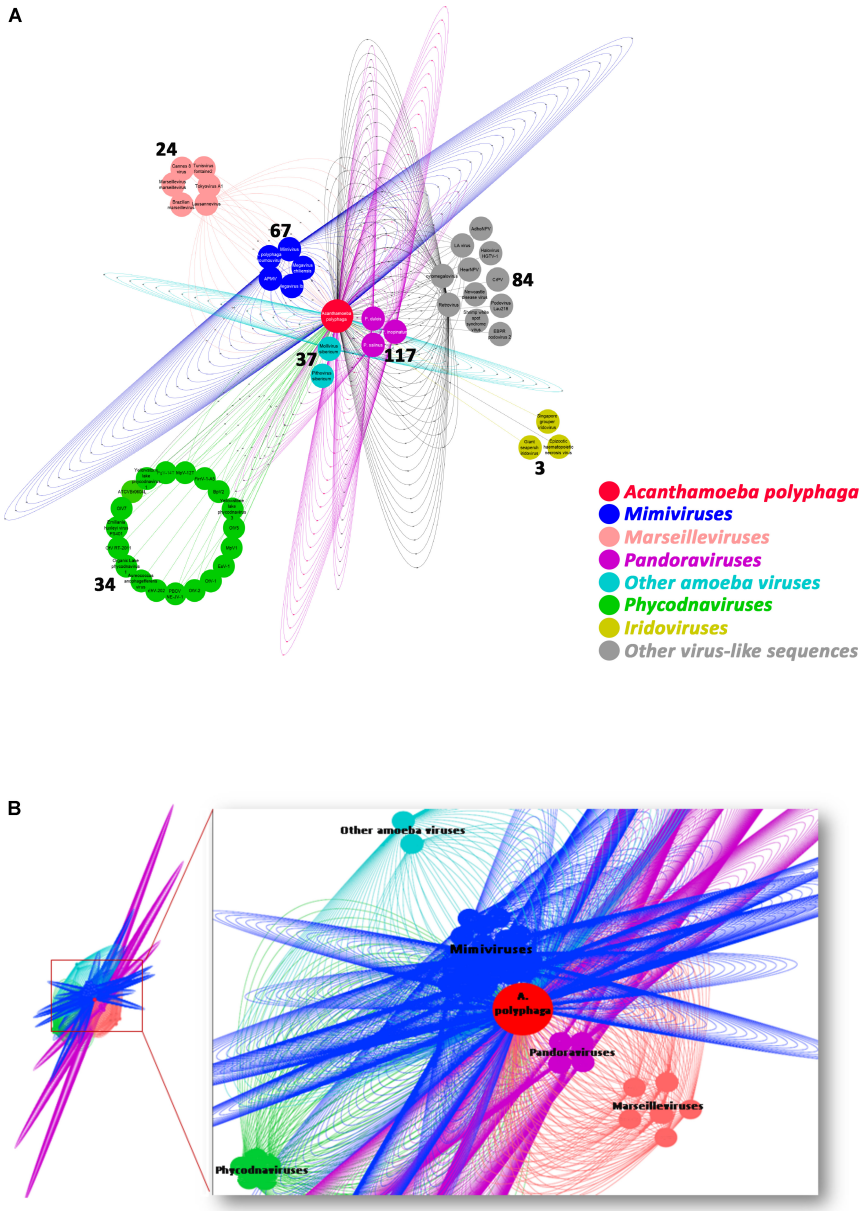


FIGURE 2 | Network of viral genes in *A. polyphaga* **(A)** and gene trafficking between *A. polyphaga* and giant viruses **(B)**. **(A)** genes were classified regarding their viral families, and the number of exchanged genes was indicated for each group of viruses; **(B)** the 1,797 genes were classified with respect to their viral families, and represented by a different color: pandoraviruses in pink; mimiviruses in dark blue; marseilleviruses in light orange; light blue includes other amoeba viruses (Mollivirus sibericum and Pithovirus sibericum); and phycodnaviruses in green.

Synteny of *Acanthamoeba polyphaga* Genes With Viral Genes as Best Hits in the Draft Genome Sequences of the 16 *Acanthamoeba* spp.

The evolution of the 10 genomic regions identified among the draft genome sequences of *Acanthamoeba* spp. and carrying the highest number of genes best matching with viruses (at least 3 genes) was analyzed. This showed that the three amoebal species *A. polyphaga* ATCC30872, *A. castellanii* ATCC 50370 and *A. pearcei* all conserved the 10 scaffolds that harbored in totality 46 viral genes. In addition, this analysis revealed a shared synteny with similar contents and co-localization of the viral-related genes in these three *Acanthamoeba* species. In contrast, for *A. quina, A. lugdunensis, A. mauritaniensis, A. rhysodes, A. palestinensis, A. healyi, A. lenticulata,* and *A. royreba*, the number of conserved genes was 24, 11, 10, 10, 8, 4, 4, and 3, respectively. In addition, none of these 46 genes was found in *A. culbertsoni*, *A. astronyxis* and *A. divionenesis* (**Figure 5**). The tree based on similarities between the 14 draft genome

TABLE 3 | Distribution of the genes predicted from the draft genome sequences of *A. polyphaga* with a viral gene as best hit in the functional "COG" categories.

COG functional category	Function	Number of genes
A	RNA processing and modification	1
E	Amino acid transport and metabolism	4
F	Nucleotide transport and metabolism	3
G	Carbohydrate transport and metabolism	1
H	Coenzyme transport and metabolism	3
J	Translation, ribosomal structure and biogenesis	2
K	Transcription	9
L	Replication, recombination and repair	19
M	Cell wall/membrane/envelope biogenesis	1
O	Post-translational modification, protein turnover, chaperones	13
P	Inorganic ion transport and metabolism	7
R	General function prediction only	16
S	Function unknown	5
T	Signal transduction	17
U	Intracellular trafficking, secretion, and vesicular transport	2
V	Defense mechanisms	1

COG, cluster of orthologous groups of proteins.

sequences and the genomes of *A. castellanii* Neff and *A. polyphaga* Linc-AP1 showed that the pattern of conservation of these genes homologous with viral genes in the draft genome sequences of the different *Acanthamoeba* species was congruent with the phylogenomic analyses since they displayed the same distribution into the groups of species.

DISCUSSION

We give here new insights into the possible interactions between *Acanthamoeba* species and giant viruses. Only one *A. castellanii* genome has been analyzed so far (Clarke et al., 2013). Our analyses primarily focused on *A. polyphaga*, for which the body of data regarding the isolation and propagation by culture of giant viruses is the greatest among amoebal species (along with that of *A. castellanii*). Great similarities were suggested between *A. polyphaga* and *A. castellanii* Neff, as *A. polyphaga* was found to harbor the majority of the 14,974 genes previously predicted through genome and transcriptome sequencing in *A. castellanii* Neff (Clarke et al., 2013). In contrast, it is worthy to note that *A. castellanii* Neff and *A. castellanii* ATCC 50370, whose draft genome sequences were analyzed here, did not cluster together based on our phylogenetic analyses. Furthermore, there is possibly an incorrect identification of sequences presented as originating from ATCC 30872 but belonging to *A. polyphaga* species. Indeed, the 18S ribosomal DNA previously described (AY026244) and part of the draft genome sequences analyzed here were not the closest among sequences from different *Acanthamoeba* species, and were not clustered in the phylogenetic analysis. These issues deserve clarification in future studies.

More than half of the genes predicted from the *A. polyphaga* draft genome sequences analyzed here had no homologs in the NCBI non-redundant sequence database. This suggests that a large part of the genome sequence of this amoeba remains unknown. Yet, a significant proportion of the predicted ORFs contained less than 100 codons. They should be considered carefully, especially when identified as ORFans or annotated as hypothetical proteins. Among annotated genes, the vast majority had homologs in amoebae, and 3% had homologs in other eukaryotes or in bacteria. Finally, a small proportion of annotated genes was mostly related to viral sequences, among which a majority belonged to the three pandoraviruses described to date and to *Mollivirus sibericum*.

The presence of sequences homologous with coding sequences of giant viruses of amoebae in eukaryotic genomes has been described by several teams (Filée, 2014; Maumus et al., 2014; Blanc et al., 2015; Sharma et al., 2016; Maumus and Blanc, 2016; Gallot-Lavallée and Blanc, 2017). Notably, MCPs were recently unexpectedly described in some *Acanthamoeba* species (Clarke et al., 2013; Maumus and Blanc, 2016; Gallot-Lavallée and Blanc, 2017). Congruently, our analysis showed that these sequences comprised two groups: firstly, the sequence of the "iridovirus-like" MCP of *A. castellanii* Neff and a sequence of Mollivirus sibericum, and, secondly, the other sequences which are related to those of a phycodnavirus (genus *Raphidovirus*), Heterosigma akashiwo virus 01. The G+C content was found to be homogeneous along the *Acanthamoeba* scaffolds carrying these genes best matching MCPs. However, this does not completely rule out the possibility that these sequences are the result of an exchange of sequences between these organisms because the GC% in some giant viruses is relatively close to that of their host, for example for pandoraviruses and *A. castellanii* (60.6% vs. 58.3%) (Antwerpen et al., 2015). These results also confirm previous observations concerning the presence of homologous sequences of MCPs in the genome of some eukaryotes (Maumus et al., 2014). Thus, these data suggest nucleotide sequence transfers between giant viruses and *Acanthamoeba,* and raise the question of the significance of homologies between genes present in giant virus and *Acanthamoeba* genomes. One hypothesis is that capsid proteins may be involved in defense mechanisms. Interestingly, it has been described that, in addition to their function as capsid proteins, the MCPs of totivirus, a RNA virus, was also able to inactivate the host mRNAs by eliminating their 5'-cap (Koonin, 2010).

Preliminary analysis of possible nucleotide sequence transfers within the *A. polyphaga* genome showed that there are 366 genes that could have been exchanged between this amoeba and viruses. We compared this gene set with the set of 267 genes recently described in *A. castellanii* Neff as putatively exchanged with viruses (Maumus and Blanc, 2016). We found that among these 366 *A. polyphaga* genes, only 30 (8.2%) were

FIGURE 3 | Continued

FIGURE 3 | Phylogenetic trees for six examples of putative viral proteins in *A. polyphaga*. Gene sequence transfer was inferred from the comparison between annotated sequences with a putative viral origin and their best hits as well as their homologs in the other *Acanthamoeba* genomes. The way of the transfer was supposed to be from giant viruses to amoebae **(a–c)** or from amoebae to giant viruses **(d–f)**. Trees are unrooted. In red: *A. polyphaga* gene; in blue: the viral homolog; in orange homologs from other *Acanthamoeba* species; in black: homologs from other organisms.

shared with the 267 genes identified in *A. castellanii* Neff as putatively exchanged with viruses. In addition, only 39 (11.0%) among the 356 genes which had as best hit a coding sequence of viruses in the draft genome sequences of *A. castellanii* ATCC 50370 were shared with the 267 genes described by Maumus and Blanc (Maumus and Blanc, 2016). A total of 11% of the 366 genes were matching with genes transcribed in *A. castellanii* Neff (Clarke et al., 2013), suggesting that they might be transcribed and play yet unknown roles. Some of these differences might be explained by differences between the sets of giant virus genomes available at the time when the different analyses were performed, as giant virus diversity expands considerably (Maumus and Blanc, 2016; Colson et al., 2017b). In addition, differences in parameters used for gene prediction and annotation and BLAST searches are likely to generate discrepant results. It should be also taken into account that the amoebal genomes analyzed here were non-assembled draft genome sequences. Among giant viruses which infect *Acanthamoeba* spp., pandoraviruses were found to share the highest number of genes with *A. polyphaga*, before mimiviruses and marseilleviruses, with 117, 65 and 24 putatively exchanged genes, respectively. This might suggests a co-evolution of *A. polyphaga* with pandoraviruses, but it is worth considering that this *Acanthamoeba* species was far less permissive to pandoraviruses than *A. castellanii* (Dornas et al., 2015). In addition, there is little evidence of gene exchanges with other viruses, including phycodnaviruses. The sense of

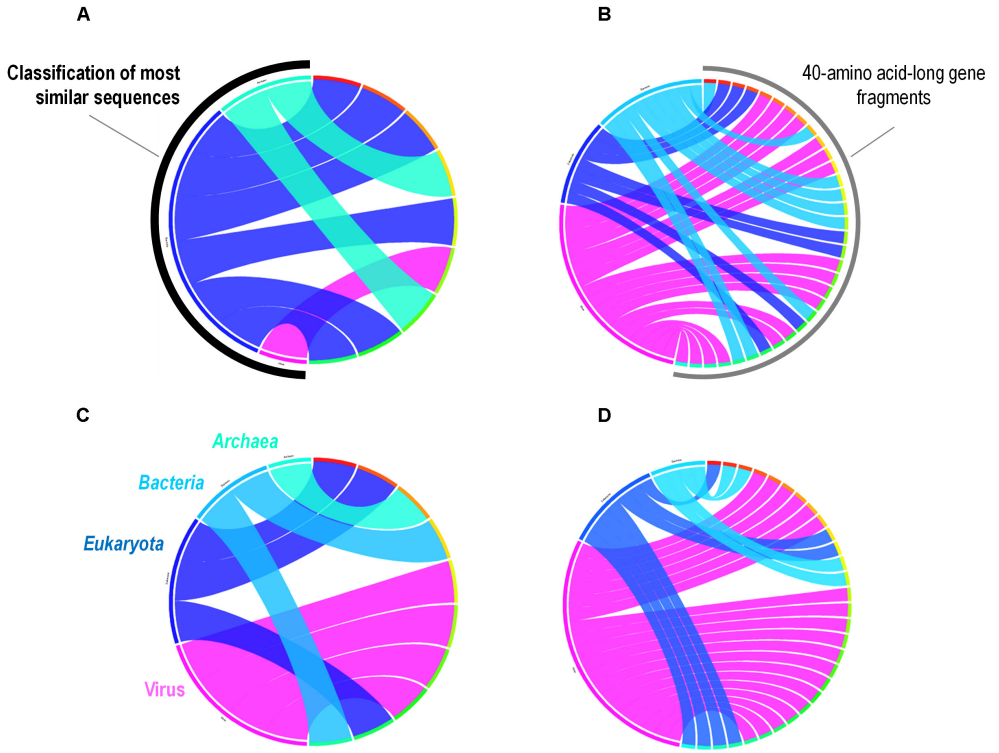


FIGURE 4 | Rhizomes of *A. polyphaga* gene fragments illustrative of the mosaicism of the genes. Forty amino acid-long fragments of four *A. polyphaga* genes either supposed to have been transferred from viruses to amoebae **(A,B)** or in the opposite way **(C,D)** were linked to their most similar sequences in the NCBI GenBank protein sequence database according to the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi), classified according to their belonging to viruses, eukaryotes, bacteria or archaea, and integrated in a circular gene data visualization. The figures were performed using the CIRCOS online tool (http://mkweb.bcgsc.ca/tableviewer/visualize/).

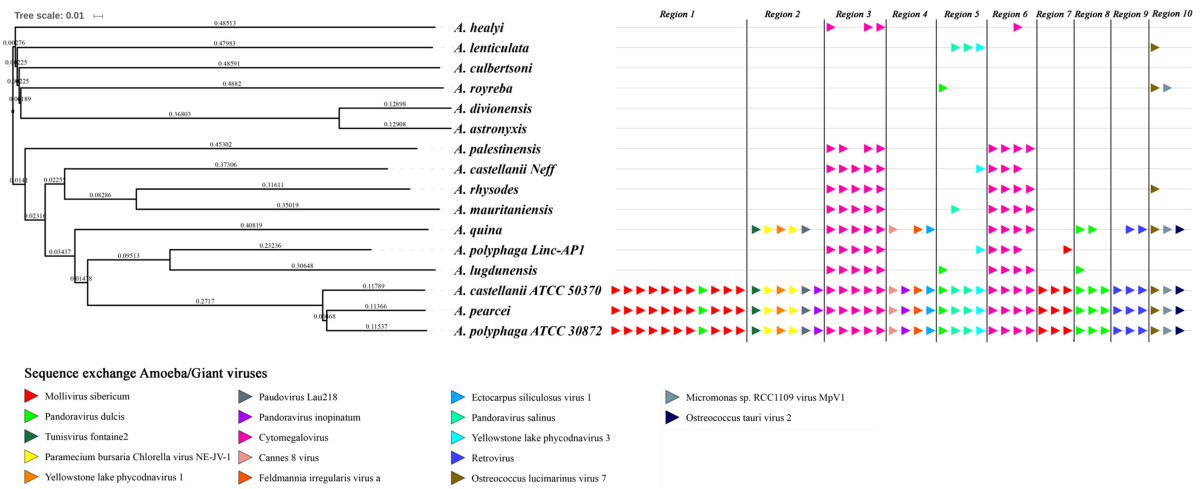


FIGURE 5 | Comparison between gene synteny for genes with giant virus genes as best match in the genome sequence of 16 *Acanthamoeba* isolates classified in 14 species, and genome tree built for these *Acanthamoeba* species. Phylogenetic tree using the complete draft genome sequences of the 16 *Acanthamoeba* strains was represented aside the synteny distribution of genes best matching with giant virus genes in 10 selected genomic regions from *Acanthamoeba* spp. Phylogenetic reconstructions were performed using the alignment of 16 draft genome sequences of *Acanthamoeba* strains classified in 14 species including two *A. polyphaga* strains and two *A. castellanii* strains, by using the progressive Mauve program (Darling et al., 2010). Pointing triangles on the right part of the Figure correspond to viral genes; viruses are represented by different colors.

the gene nucleotide sequence transfers remained undetermined in a large majority of cases. This is due to an insufficient number of matches, or an absence of significantly delineated cluster in the phylogenetic trees. Nevertheless, for some of the phylogenies, a transfer from *Acanthamoeba* to giant viruses was the most parsimonious evolutionary hypothesis, indicating that giant viruses are not only bags of genes but contribute to the gene sequence flow. Estimating times of divergence might be helpful in more extensive analyses that would be conducted on giant viruses of amoebae and several *Acanthamoeba* species in order to try inferring the sense of gene sequence transfers between giant viruses and their amoebal hosts. Besides, the analyses of fragments of *A. polyphaga* genes that had a giant virus sequence as best hit showed a different, more complex pattern of best hits compared to the analysis of the whole genes, with mosaics of sequences from eukaryotes, bacteria, archaea and giant viruses as best matches. A similar pattern has also been illustrated by fragmenting a gene encoding an aminoacyl-tRNA synthetase in Klosneuvirus, a mimivirus relative (Abrahao et al., 2018). These findings may extend to genes the paradigm that no single tree can define the mosaic origin of genomes, which may result from nucleotide sequence transfers rather than from gene transfers (Dagan and Martin, 2006; Merhej et al., 2011). These mosaic patterns that affect the sequences of genes may represent a pitfall for the robustness of phylogenetic analyses and inference of the putative way of nucleotide sequence transfer. This further suggests that the vision of species in Darwin's tree of life is rather outdated, and that each organism has a complex family tree that is a testimony of its chimerical origin and is better represented in the form of a rhizome than of a tree (Raoult, 2010; Merhej et al., 2011). Overall, the global analysis of all giant virus homologs to *A. polyphaga* predicted genes potentially demonstrates the complexity of the putative gene trafficking between this amoeba and giant viruses, with 1,797 genes involved, although only 366 *A. polyphaga* genes have viral genes as best matches. This might further suggest intermediate interactions with organisms other than viruses. Moreover, these results highlight the fact that the gene flow was not a one way mechanism and likely resulted from the sympatric lifestyle of giant viruses in amoebae (Moliner et al., 2010).

The comparisons of possible gene sequence transfers between the two species *A. polyphaga* and *A. castellanii* ATCC 50370 and some representatives of giant viruses (pandoraviruses, *Acanthamoeba polyphaga* mimivirus and Marseillevirus) shows that the differences regarding the number and the nature of potentially transferred genes remain limited within the same viral family. However, the number and the nature of potentially transferred genes were found to be more variable when considering different families of giant viruses. These observations might be explained by differences in the frequency of interactions between some *Acanthamoeba* species and some giant viruses, and between genes used by both the amoebae and the giant viruses to deal with these interactions. The recent study by Clarke et al. (2013) showed that the plasticity of protists living in community with microbes is as important as that of bacteria with the same lifestyle. This supports the hypothesis that an essential explanation of the chimerism of the genomes of organisms and micro-organisms that live sympatrically relies on this lifestyle, and possibly on their genomic plasticity according to their phylogenetic origin (eukaryotic, bacterial, archaeal or viral). These observations may also reflect the time of sequence exchange through the evolutionary course of both giant viruses and amoeba species. Phylogenetic and synteny analyses of viral genes performed in the present study suggest that the sequence exchanges between *Acanthamoeba* species and giant viruses occurred along the *Acanthamoeba* speciation events and evolved through species-specific events. Among the *Acanthamoeba* species, *A. castellanii* ATCC 50370 and *A. pearcei* were found to be the most structurally conserved with *A. polyphaga* regarding their content of genes homologous with viral genes. These results did follow the expectation that decreased phylogenetic distance would correspond to increased levels of genome preservation.

In summary, this work opens up several potential avenues for future works on the interactions between *Acanthamoeba* species and giant viruses. The annotation of all *Acanthamoeba* species and their accurate identification are important tasks for a greater understanding of why some amoebae are more susceptible than others to giant viruses, and possibly to other microorganisms. Thus, the considerable diversity of gene repertoires among *Acanthamoeba* species might lead to differences regarding potential interactions with giant viruses. The characterization of the genes present and absent in the different species of *Acanthamoeba* could be performed and correlated with the observed phenotypic differences that need to be studied more extensively.

AUTHOR CONTRIBUTIONS

PC and BLS designed the study. NC, PC, and AL performed the experiments. NC, PC, AL, PP, DR, and BLS analyzed the data. All the authors contributed to manuscript redaction and review.

FUNDING

This work was supported by a grant from the French State managed by the National Research Agency under the "Investissements d'avenir (Investments for the Future)" program with the reference ANR-10-IAHU-03 (Méditerranée Infection). NC was financially supported through a grant from the Infectiopole Sud foundation. This work was supported by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02098/full#supplementary-material

REFERENCES

Abrahao, J., Silva, L., Santos Silva, L., Bou Khalil, J. Y., Rodrigues, R., Arantes, T., et al. (2018). Tailed giant tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* 9:749. doi: 10.1038/s41467-018-03168-3161

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-80362

Antwerpen, M. H., Georgi, E., Zoeller, L., Woelfel, R., Stoecker, K., and Scheid, P. (2015). Whole-genome sequencing of a *Pandoravirus* isolated from keratitis-inducing *Acanthamoeba*. *Genome Announc.* 3:e00136-15. doi: 10.1128/genomeA.00136-15

Bertelli, C., and Greub, G. (2012). Lateral gene exchanges shape the genomes of amoeba-resisting microorganisms. *Front. Cell. Infect. Microbiol.* 2:110. doi: 10.3389/fcimb.2012.00110

Blanc, G., Gallot-Lavallée, L., and Maumus, F. (2015). Provirophages in the *Bigelowiella* genome bear testimony to past encounters with giant viruses. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5318–E5326. doi: 10.1073/pnas.1506469112

Boyer, M., Azza, S., Barrassi, L., Klose, T., Campocasso, A., Pagnier, I., et al. (2011). Mimivirus shows dramatic genome reduction after intraamoebal culture. *Proc. Natl. Acad. Sci.* 108, 10296–10301. doi: 10.1073/pnas.1101118108

Clarke, M., Lohan, A. J., Liu, B., Lagkouvardos, I., Roy, S., Zafar, N., et al. (2013). Genome of *Acanthamoeba castellanii* highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling. *Genome Biol.* 14:R11. doi: 10.1186/gb-2013-14-2-r11

Colson, P., De Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D. K., et al. (2013). "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* 158, 2517–2521. doi: 10.1007/s00705-013-1768-1766

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

Colson, P., La Scola, B., and Raoult, D. (2017b). Giant viruses of amoebae: a journey through innovative research and paradigm changes. *Annu. Rev. Virol.* 4, 61–85. doi: 10.1146/annurev-virology-101416-141816

Dagan, T., and Martin, W. (2006). The tree of one percent. *Genome Biol.* 7:118. doi: 10.1186/gb-2006-7-10-118

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

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

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

Filée, J. (2014). Multiple occurrences of giant virus core genes acquired by eukaryotic genomes: the visible part of the iceberg? *Virology* 466–467, 53–59. doi: 10.1016/j.virol.2014.06.004

Filée, J., Siguier, P., and Chandler, M. (2007). I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends Genet.* 1, 10–15. doi: 10.1016/j.tig.2006.11.002

Gallot-Lavallée, L., and Blanc, G. (2017). A glimpse of nucleo-cytoplasmic large DNA virus biodiversity through the eukaryotic genomics window. *Viruses* 9:17. doi: 10.3390/v9010017

Greub, G., and Raoult, D. (2004). Microorganisms resistant to free-living amoebae. *Clin. Microbiol. Rev.* 17, 413–433. doi: 10.1128/CMR.17.2.413

Huson, D. H., Beier, S., Flade, I., Górska, A., and El-hadidi, M. (2016). MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12:e1004957. doi: 10.1371/journal.pcbi.1004957

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

Khalil, J. Y. B., Andreani, J., and La Scola, B. (2016). Updating strategies for isolating and discovering giant viruses. *Curr. Opin. Microbiol.* 31, 80–87. doi: 10.1016/j.mib.2016.03.004

Koonin, E. V. (2010). Taming of the shrewd: novel eukaryotic genes from RNA viruses. *BMC Biol.* 8:2. doi: 10.1186/1741-7007-8-2

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

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

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (2015). In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting *Acanthamoeba*. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5327–E5335. doi: 10.1073/pnas.1510795112

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., and Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* 33, 6494–6506. doi: 10.1093/nar/gki937

Marciano-cabral, F., and Cabral, G. (2003). *Acanthamoeba* spp. as agents of disease in humans. *Clin. Microbiol. Rev.* 16, 273–307. doi: 10.1128/CMR.16.2.273

Maumus, F., and Blanc, G. (2016). Study of gene trafficking between *Acanthamoeba* and giant viruses suggests an undiscovered family of amoeba-infecting viruses. *Genome Biol. Evol.* 8, 3351–3363. doi: 10.1093/gbe/evw260

Maumus, F., Epert, A., Nogué, F., and Blanc, G. (2014). Plant genomes enclose footprints of past infections by giant virus relatives. *Nat. Commun.* 5:4268. doi: 10.1038/ncomms5268

Merhej, V., Notredame, C., Royer-Carenzi, M., Pontarotti, P., and Raoult, D. (2011). The rhizome of life: the sympatric *Rickettsia felis* paradigm demonstrates the random transfer of DNA sequences. *Mol. Biol. Evol.* 28, 3213–3223. doi: 10.1093/molbev/msr239

Moliner, C., Fournier, P. E., and Raoult, D. (2010). Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiol. Rev.* 34, 281–294. doi: 10.1111/j.1574-6976.2009.00209.x

Moreira, D., and Brochier-Armanet, C. (2008). Giant viruses, giant chimeras: the multiple evolutionary histories of mimivirus genes. *BMC Evol. Biol.* 8:12. doi: 10.1186/1471-2148-8-12

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

Raoult, D. (2010). The post-Darwinist rhizome of life. *Lancet* 375, 104–105. doi: 10.1016/S0140-6736(09)61958-61959

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004). The 1.2-megabase genome sequence of Mimivirus. *Science* 306, 1344–1350. doi: 10.1126/science.1101485

Raoult, D., and Boyer, M. (2010). Amoebae as genitors and reservoirs of giant viruses. *Intervirology* 53, 321–329. doi: 10.1159/000312917

Reteno, D. G., Benamar, S., Khalil, J. B., Andreani, J., Armstrong, N., Klose, T., et al. (2015). Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* 89, 6585–6594. doi: 10.1128/JVI.00115-115

Rodríguez-Zaragoza, S. (1994). Ecology of free-living amoebae. *Crit. Rev. Microbiol.* 20, 225–241. doi: 10.3109/10408419409114556

Salah, I. B., Ghigo, E., and Drancourt, M. (2009). Free-living amoebae, a training field for macrophage resistance of mycobacteria. *Clin. Microbiol. Infect.* 15, 894–905. doi: 10.1111/j.1469-0691.2009.03011.x

Sharma, V., Colson, P., Pontarotti, P., and Raoult, D. (2016). Mimivirus inaugurated in the 21st century the beginning of a reclassification of viruses. *Curr. Opin. Microbiol.* 31, 16–24. doi: 10.1016/j.mib.2015.12.010

Sherr, E. B., and Sherr, B. F. (2002). Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek* 81, 293–308. doi: 10.1023/A:1020591307260

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432. doi: 10.1093/bioinformatics/btq675

Tatusov, R. L. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33

Yutin, N., Wolf, Y. I., Raoult, D., and Koonin, E. V (2009). Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol. J.* 6:223. doi: 10.1186/1743-422X-6-223

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 22 December 2017
doi: 10.3389/fmicb.2017.02562

Genome Characterization of the First Mimiviruses of Lineage C Isolated in Brazil

Felipe L. Assis[1], Ana P. M. Franco-Luiz[1], Raíssa N. dos Santos[2], Fabrício S. Campos[3], Fábio P. Dornas[1], Paulo V. M. Borato[1], Ana C. Franco[2], Jônatas S. Abrahao[1], Philippe Colson[4] and Bernard La Scola[4]**

[1] Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [2] Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [3] College of Veterinary Medicine and Agronomy, University of Brasília, Brasília, Brazil, [4] CNRS 7278, IRD 198, INSERM 1095, UM63, IHU – Méditerranée Infection, AP-HM, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Aix-Marseille Université, Marseille, France

The family *Mimiviridae*, comprised by giant DNA viruses, has been increasingly studied since the isolation of the *Acanthamoeba polyphaga mimivirus* (APMV), in 2003. In this work, we describe the genome analysis of two new mimiviruses, each isolated from a distinct Brazilian environment. Furthermore, for the first time, we are reporting the genomic characterization of mimiviruses of group C in Brazil (Br-mimiC), where a predominance of mimiviruses from group A has been previously reported. The genomes of the Br-mimiC isolates Mimivirus gilmour (MVGM) and Mimivirus golden (MVGD) are composed of double-stranded DNA molecules of ∼1.2 Mb, each encoding more than 1,100 open reading frames. Genome functional annotations highlighted the presence of mimivirus group C hallmark genes, such as the set of seven aminoacyl-tRNA synthetases. However, the set of tRNA encoded by the Br-mimiC was distinct from those of other group C mimiviruses. Differences could also be observed in a genome synteny analysis, which demonstrated the presence of inversions and loci translocations at both extremities of Br-mimiC genomes. Both phylogenetic and phyletic analyses corroborate previous results, undoubtedly grouping the new Brazilian isolates into mimivirus group C. Finally, an updated pan-genome analysis of genus *Mimivirus* was performed including all new genomes available until the present moment. This last analysis showed a slight increase in the number of clusters of orthologous groups of proteins among mimiviruses of group A, with a larger increase after addition of sequences from mimiviruses of groups B and C, as well as a plateau tendency after the inclusion of the last four mimiviruses of group C, including the Br-mimiC isolates. Future prospective studies will help us to understand the genetic diversity among mimiviruses.

Keywords: ***Mimiviridae*, pan-genome, genomics, giant virus, mimivirus**

OPEN ACCESS

Edited by:
Steven M. Short, University of Toronto Mississauga, Canada

Reviewed by:
Masaharu Takemura, Tokyo University of Science, Japan
Yael Mutsafi Benhalevy, National Institutes of Health (NIH), United States

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

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

Received: *27 October 2017*
Accepted: *11 December 2017*
Published: *22 December 2017*

Citation:
Assis FL, Franco-Luiz APM, dos Santos RN, Campos FS, Dornas FP, Borato PVM, Franco AC, Abrahao JS, Colson P and La Scola B (2017) Genome Characterization of the First Mimiviruses of Lineage C Isolated in Brazil. Front. Microbiol. 8:2562. doi: 10.3389/fmicb.2017.02562

INTRODUCTION

Since the serendipitous discovery of *Acanthamoeba polyphaga mimivirus* (APMV) in 2003, dozens of studies have been conducted to describe how widespread and diverse this new viral family is La Scola et al. (2003, 2010), Raoult et al. (2004), Arslan et al. (2011), Colson et al. (2011a,b), Legendre et al. (2012), Boughalmi et al. (2013b,c), Saadi et al. (2013a), Yoosuf et al. (2014a,b).

Concomitantly, researchers have been working on the biology and molecular characterization of other mimivirus relatives isolated from several human and environmental samples, the latter of which include cooling water tower, freshwater, saltwater, soil, leech, oyster, and other sources collected in many countries in Oceania, Europe, Asia, Africa, and South America (La Scola et al., 2008; Fischer et al., 2010; Arslan et al., 2011; Yoosuf et al., 2012; Boughalmi et al., 2013a,b,c; Pagnier et al., 2013; Saadi et al., 2013a,b; Campos et al., 2014; Bajrai et al., 2016; Takemura et al., 2016). During those studies, notable sets of genes seemingly encoded by the genome of these new viruses were observed. These included genes encoding tRNA sequences, aminoacyl-tRNA synthetases, and peptide synthesis factors. Equally surprising was that mimiviruses can be associated with small viruses, which were named virophages in analogy to bacteriophages that infect bacterial hosts (La Scola et al., 2008). Some mimiviruses were recently predicted to encode a defense system named the MIMIVIRE, which enables them to target specific virophages (Levasseur et al., 2016). However, all these astonishing discoveries could be the "tip of the iceberg" regarding mimivirus features, as ∼50% of the sequences of these viruses encode proteins that are hypothetical, i.e., without a defined function (La Scola et al., 2003; Raoult et al., 2004).

The mimiviruses have a semi-icosahedral 410–550 nm in diameter capsid, with a symmetry breaking at a single vertex of the particle forming a five-branch star structure, called the 'stargate' (Zauberman et al., 2008). The capsid surface is covered, except at the "stargate" vertex, by a 150-nm thick fibril layer, involved in a matrix with a composition initially thought to be similar to peptidoglycan. Although the mimiviruses have been isolated using co-culture on amoebae of the genus *Acanthamoeba*, knowledge about their natural reservoir as well as their host range is still limited. The mimiviruses replicate in the host cytoplasm in a replication factory that is formed after the genome is released (Suzan-Monti et al., 2007; Mutsafi et al., 2010; Colson et al., 2017). The genomes of mimiviruses are comprised by a linear dsDNA molecule that is 0.92–1.22 Mb long and encodes 930–1,178 proteins (Raoult et al., 2004). The genome of the prototype Mimivirus was described to present two inverted repeats of about 900 nucleotides near both extremities, suggesting that the Mimivirus genome might adopt a circular topology during viral replication, as in some other NCLDVs (Raoult et al., 2004).

The family *Mimiviridae* is comprised by two genera, named (1) *Mimivirus*, composed of mimiviruses infecting amoebal species, and (2) *Cafeteriavirus*, a distantly related mimiviruses group comprised by the type species *Cafeteria roenbergensis virus* (CroV; which infects a marine heterotrophic bi-flagellate) International Committee on Taxonomy of Viruses [ICTV], 2017. Other related distant mimiviruses have been associated with CroV, including Organic lake phycodnaviruses and *Phaeocystis globosa* virus (Yutin et al., 2013). The recently described klosneuviruses also seem to be related to *Mimiviridae* members (Schulz et al., 2017). The genus *Mimivirus* can be divided into three lineages A, B, and C, according to phylogenomic data including phylogenies based on conserved core genes, for example family B DNA polymerase and ribonucleotide reductase encoding genes (Boyer et al., 2010; Colson et al., 2012; Legendre et al., 2012; Campos et al., 2014).

We isolated the first Brazilian mimivirus strain, named Samba virus (SBV), from a water sample collected in the Amazon region in 2011. Phylogenomic analyses clustered the SBV into mimivirus lineage A (Campos et al., 2014), which includes the APMV, the prototype species of family *Mimiviridae*. More recently, Brazilian mimivirus strains have been isolated and/or detected from fresh water, oyster, sewage, humans, and both wild and domestic mammals, and their biological and molecular characterization have been reported (Dornas et al., 2014, 2016, 2017; Andrade et al., 2015; Boratto et al., 2015). Curiously, all Brazilian mimivirus strains were classified into mimivirus lineage A, suggesting that this lineage is the most widespread in Brazil (Andrade et al., 2015; Assis et al., 2015; Boratto et al., 2015). In addition, Assis et al. (2015) described the pan-genome of mimivirus lineage A, which was composed of 1129 clusters of orthologous groups (COGs) of proteins encoded by all genomes available at that time. All these data led us to ask more questions about the diversity of mimiviruses circulating in Brazil and resulted in the decision to conduct additional prospective studies. In this way, Dornas et al. (2015), using a panel of protozoa (*Acanthamoeba castellanii* [AC], *Acanthamoeba polyphaga* [AP], *Acanthamoeba griffinii* [AG] and *Vermamoeba vermiformis* [VV]), were able to isolate 62 new mimivirus-like strains from sewage, sludge, water, wet soil, and lake sediment collected from different areas of the Pampulha lagoon in Belo Horizonte, Minas Gerais, Brazil (Dornas et al., 2015). A higher prevalence of lineage A mimiviruses (90.3%) was observed, followed by lineage C mimiviruses (6.4%) and lastly lineage B mimiviruses (3.2%). However, neither further analysis of the biological and molecular features of these viruses nor phylogenies were provided, once the classification of these new isolates into lineages was inferred based on BLAST hits obtained against the NCBI nt database.

In this work, we report the molecular and phylogenetic analysis of two Brazilian mimiviruses from lineage C (Br-mimiC): (1) Mimivirus gilmour (MVGM) – isolated from water collected at Pampulha lagoon by Dornas et al. (2015); (2) Mimivirus golden (MVGD) – isolated from golden mussels (*Limnoperna fortunei*) collected from Guaíba Lake, Rio Grande do Sul, Brazil, in July 2014. Both Br-mimiC viruses were isolated using the protozoa AP as support for co-culture. In addition, we conducted an updated pan-genome analysis of all available genomes of mimiviruses from lineages A to C.

MATERIALS AND METHODS

Sample Collection

A collaborative effort involving the Aix-Marseille University (France), and the Federal Universities of Minas Gerais and Rio Grande do Sul (Brazil) was established aiming to conduct prospective studies of giant viruses in different regions and environments in Brazil. All collection procedures were performed with the authorization of IBAMA-SISBIO (number 34293-2). For

this work, water samples were collected in sterile tubes from Pampulha lagoon in September 2014, and were directly used for inoculation procedures. In addition, golden mussels (*L. fortunei*) were collected from Guaíba Lake in July 2014 (30°01′59″ S, 51°13′48″ W) (Dos Santos et al., 2016). The mussels were collected from the lake bottom at a depth of 2 m, and they were attached to a metal grid that had been submerged for 6 months before the date of collection. Golden mussels were submerged for 15 min in 70% ethanol for superficial shell decontamination. Subsequently, the valves were opened and the inner water was collected and diluted in 1 mL of saline buffer (PBS). The samples were pooled, totaling eight pools. These pools were homogenized with 1 mL of PBS, centrifuged at 10,000 × *g*, filtered through a 0.45-μm pore-size membrane. The resulting filtrate was treated with 10 U/μL of Penicillin-GIBCO by Life Technologies to prevent bacterial contamination. All the samples were stored at 4°C until the inoculation procedures.

Virus Isolation

The MVGM sample was isolated using co-culture of AP strain LINC AP1 previously cultured in a 75-cm^2 cell culture flask with 30 ml of peptone-yeast extract-glucose medium (PYG) at 30°C for 25 h. The culture supernatant was pelleted by centrifugation, suspended in PAS supplemented with an antibiotic mix containing 10 μL of ciprofloxacin (4 μg/mL; Panpharma, Z.I., Clairay, France), 10 μL of vancomycin (4 μg/mL; Mylan, Saint-Priest, France), 10 μL of colimycin (500 IU/mL; Sanofi Aventis, Paris, France), 10 μL of rifampicin (4 μg/mL; Sanofi Aventis), and 10 μL of fungizone (100 μg/mL; Bristol-Myers Squibb, Rueil-Malmaison, France), and dispensed in 0.5 ml amounts to the wells of a 24-well plate with a suspension cell concentration of 10^6 cells/ml. After that, 100 μL of samples were inoculated into wells and incubated at 30°C for 4 days. The sub-cultures were performed twice on fresh amoebae in a 1-10th dilution. A negative amoebal control was used in each microplate (Dornas et al., 2015). For the MVGD strain, the amoeba support for co-culture were AP genotype T4 previously cultured in 10 mL of Peptona-Yeast Extract-Glucose (PYG) medium at 30°C in 25-cmł culture flasks supplemented with 50 μg of gentamicin. After 48 h, the cells were harvested and centrifuged. The pellet was re-suspended in sterile PAS (Page's amoeba saline), and 10^4 amoebas per well were cultured in 24-well microplates. After 24 h, 100 μL of samples were inoculated into wells and incubated at 30°C for 3 days. The sub-cultures were performed as previously mentioned, and amoeba cells were assessed daily for the presence of viruses and for cytopathic effects on the cell monolayer.

DNA Extraction and Genome Sequencing

For MVGM strain, viral DNA was extracted with the automated EZ1 Virus Mini-Kit v.2 kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. DNA quality and concentration were checked, using a nanodrop spectrophotometer (Thermo Scientific, Waltham, MA, United States). For the MVGD strain, the supernatant of the *A. polyphaga*-infected cells was collected, and centrifuged at 5,000 × *g* for 5 min. The cell-free virus particles were pelleted on a 25% sucrose cushion by ultracentrifugation (Sorvall Combi) at 33,000 × *g* for 2 h at 4°C. The pellet was re-suspended in Tris-EDTA-NaCl buffer (TEN). In order to remove the nucleic acids not protected by the capsid, the preparation was treated with 100 U of DNAse I (Roche) and 100 U of RNAse (Invitrogen) at 37°C for 1 h. Next, the virus DNA was extracted using phenol–chloroform (Sambrook and Russell, 2001) and re-suspended in ultrapure water. The quality and amount of virus DNA was analyzed using a NanoSpec and Qubit apparatus (Life Technologies). Both extracted viral DNA were submitted to sequencing performed in a MiSeq (Illumina) apparatus with paired-end applications (2 bp × 150 bp). The pair-end samples were prepared with a Nextera XT DNA sample prep kit.

Genome Assembly and Annotation

After sequencing, reads from MVGM and MVGD were *de novo* assembled using Geneious and SPADES softwares. The gene predictions were performed using RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008) and GeneMarkS (Besemer et al., 2001) tools. Transfer RNA (tRNA) sequences were identified using the tRNAscan-SE tool (Schattner et al., 2005). The functional annotations were inferred by BLAST searches against the GenBank NCBI non-redundant protein sequence database (nr) (e-value $< 1 \times 10^{-3}$), the set of COGs of the NCLDVs (named NCVOGs) (Altschul et al., 1990; Yutin et al., 2009) and by searching specialized databases through the Blast2GO platform (Conesa et al., 2005). The genome annotations were then manually revised and curated. The predicted open reading frames (ORFs) smaller than 50 amino acids (aa) and that had no hit in any database were discarded. The ORFs longer than 50 aa without hits in any database (ORFans) were kept.

Comparative Genomic and Pan-genome Analysis

The synteny among mimiviruses from distinct lineages was checked using MAUVE program (Darling et al., 2010). The OrthoMCL tool (Chen et al., 2006) was used to identify the paralog families from Br-mimiC genomes, while Proteinortho5.pl tool (Lechner et al., 2011) was used to identify orthologous gene sequences shared by Br-mimiC. The average amino acid identity (AAI) calculator tool (Rodriguez-R and Konstantinidis, 2014) was used to compare identity between orthologous genes from Br-mimiC strains and representative mimiviruses of lineages A-C. To estimate the size of the pan-genome of the family *Mimiviridae*, their predicted proteins were clustered using the Proteinortho5.pl program (Lechner et al., 2011), using an aa sequence identity of 30% and a sequence coverage of 50% as thresholds. We also described pan-genome and core genes size variation by stepwise inclusion of each new virus annotation in the pairwise comparisons of the gene contents of all available mimivirus genome sequences.

Phylogeny

The aa sequence alignments and phylogenetic trees were built using the MEGA6 software (Tamura et al., 2013) and the

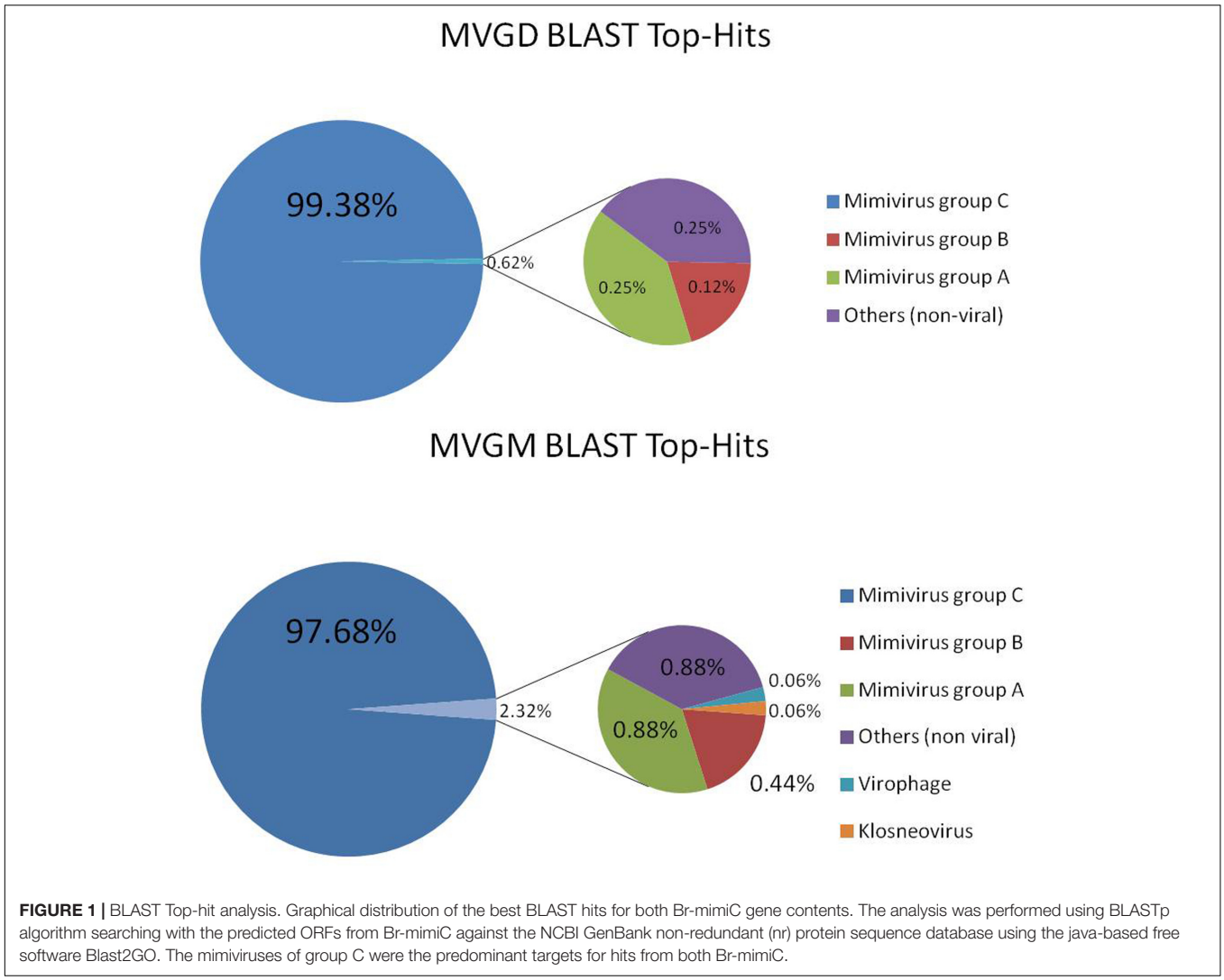


FIGURE 1 | BLAST Top-hit analysis. Graphical distribution of the best BLAST hits for both Br-mimiC gene contents. The analysis was performed using BLASTp algorithm searching with the predicted ORFs from Br-mimiC against the NCBI GenBank non-redundant (nr) protein sequence database using the java-based free software Blast2GO. The mimiviruses of group C were the predominant targets for hits from both Br-mimiC.

maximum likelihood method. Phylogenetic reconstructions were based on individual alignment of the five core genes, namely the family B DNA polymerase, the D6/D11 helicase, the VV A18 helicase, the D5 primase-helicase, and the Major Capsid Protein. In addition, we performed a hierarchical clustering based on the gene presence/absence pattern of 5443 NCVOGs, using the MeV tool (Eisen et al., 1998) with Pearson correlation as distance metric. The phylogenetic tree was visualized using the FigTree v1.4.3 tool (available online: http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS

General Features of Br-mimiC Genomes

The genomes of MVGM (GenBank number: MG602507) and MVGD (GenBank number: MG602508) are double-stranded DNA molecules of 1,258,663 base pairs (bp) (partial sequence) and 1,248,960 base pairs (complete sequence) encoding 1,135 and 1,127 ORFs, respectively. The ORFs length of both Br-mimiC ranged from 37 to 2,907 aa, with an average length of 326 aa. The BLAST analysis (coverage > 90%; identity > 80%; e-value < 10e-5) against the NCBI nr database (updated in October, 2017) identified 1088 and 1090 hits for MVGM and MVGD sequences, respectively. Furthermore, we identified 28 and 19 ORFans into MVGM and MVGD genomes, respectively. In addition, 19 and 18 ORFs without BLAST hit and smaller than 50 aa were not include in the subsequent analysis neither in the final annotation of MVGM and MVGD genomes, respectively.

The comparison between the Br-mimiC viruses genomes showed the presence of 1,042 orthologous proteins, whereas 66 and 61 proteins are unique to MVGM and MVGD, respectively. The set of unique genes of the MVGD included 18 ORFans, besides hypothetical proteins, ankyrins, F-box and FNIP repeat-containing proteins, collagen-like proteins, BTB POZ domain and WD-repeat proteins and a cholinesterase-like protein. With the exception of the cholinesterase-like protein found in the MVGD genome, the set of unique genes of the MVGM was comprised by the same classes of protein, besides a DNA primase and a putative transposase. In addition, the MVGM and MVGD

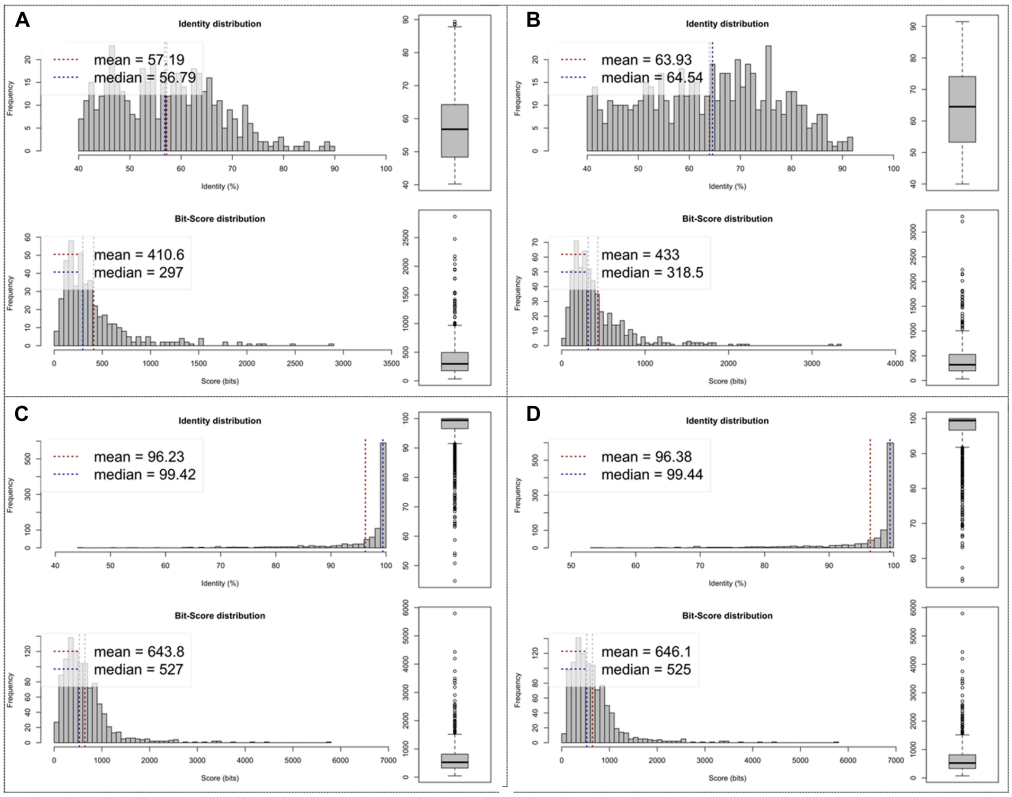


FIGURE 2 | Average amino acid identity (AAI). In this analysis, estimates were reached out using both best hits (one-way AAI) and reciprocal best hits (two-way AAI) between two datasets of proteins from the Br-mimiC isolates and representative strains from mimivírus groups A to C. Plots **(A,B)** demonstrate the AAI between MVGM and mimiviruses from group A to group B, respectively. Plot **(C)** shows the AAI between MVGM and MVGD. Plot **(D)** shows the AAI between MCHV (group C prototype strain) and MVGD.

TABLE 1 | Distribution of aminoacyl-tRNA synthetases encoded by mimivirus from group A to group C, besides Br-mimiC isolates.

	Mimivirus strains			
Aminoacyl-tRNA synthetase	**Mimivirus A**	**Mimivirus B**	**MCHV**	**Br.mimi C**
Tyrosyl	V	V	V	V
Cysteinyl	V	V	V	V
Methionyl	V	V	V	V
Arginyl	V	V	V	V
Isoleucyl	X	V	V	V
Asparaginyl	X	X	V	V
Tryptophanyl	X	X	V	V

MCHV, Megavirus chilensis; V, presence; X, absence.

genomes encoded to 551 and 558 proteins without defined function, respectively.

Both genomes showed a very similar G+C content (∼26%), genome density (∼0.890 genes per kbp), coding percentage (∼88.5%), and average gene length (995 bp). The best hit analysis for the sequences predicted in these Br-mimiC genomes showed the highest percentage (average of 98.5%) of hits against mimivirus group C sequences (**Figure 1**). The average AAI analysis (**Figure 2**) corroborated the best hit analysis showing the greatest AAI value between sequences from Br-mimiC and other mimiviruses of group C (∼96%), followed by mimiviruses of group B (63.9%) and mimiviruses of group A (57.1%). When compared between each other, the Br-mimiC showed an AAI of 96.3% (**Figure 2**). Beyond, the ORFs predicted into Br-mimiC genomes possess orthologs into other mimiviruses, hosts and/or sympatric organisms, beside virophage and other giant viruses. Furthermore, we observed the presence of seven aminoacyl (tyrosyl, cysteinyl, methionyl, arginyl, isoleucyl, asparaginyl, and tryptophanyl) tRNA synthetases (aaRS) in Br-mimiC, which has been described as a signature of mimiviruses of group C (Colson et al., 2013), while mimiviruses of groups A and B encode four and five aaRS, respectively (**Table 1**). Although the best hit analysis has shown a match against a virophage sequence, we did not detect those mimivirus-related virus associated with Br-mimiC.

Even sharing several genetic features, such as a low G+C content and large and similar genome sizes and gene repertories, the Br-mimiC isolates presented singular features which allowed distinguishing them as two distinct isolates. One of the main differences between the Br-mimiC viruses is the presence of six tRNA molecules (2x Leu-TAA, Leu-CAA, Trp-CCA, His-GTG, and Cys-GCA) encoded by MGMV, while the MGDV was

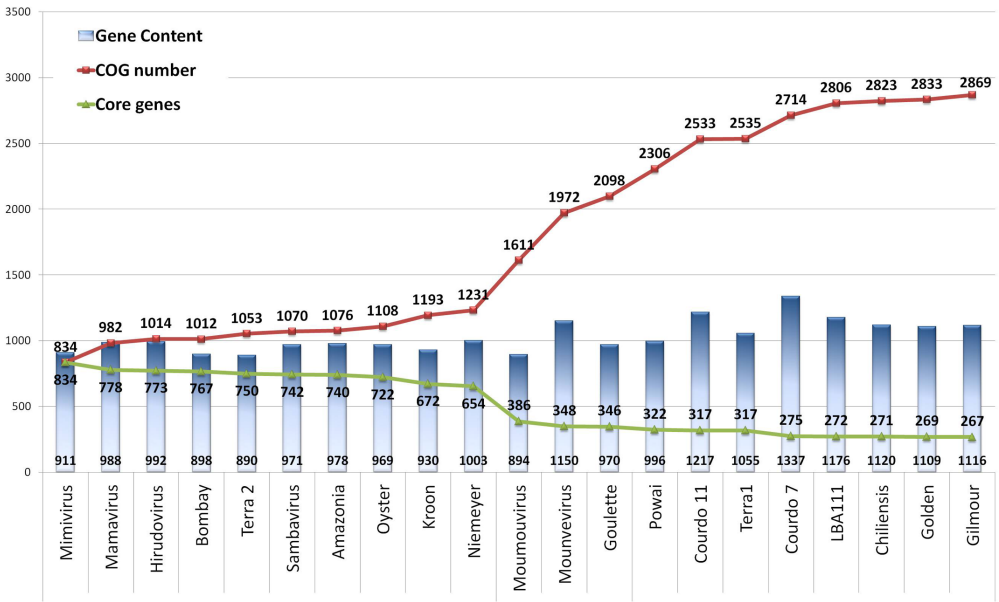


FIGURE 3 | Evolution of the pan-genome (red line) and core genome (green line) size of the family *Mimiviridae*. Numbers at the base of blues bars refer to the number of genes encoded by each virus strain. Numbers at line nodes represent the cumulative COG (red line) and core genes (green line) numbers after the inclusion of a new genome. The COGs definition was performed by using the Proteinortho5 tool, using AAI and coverage of 30 and 50%, respectively.

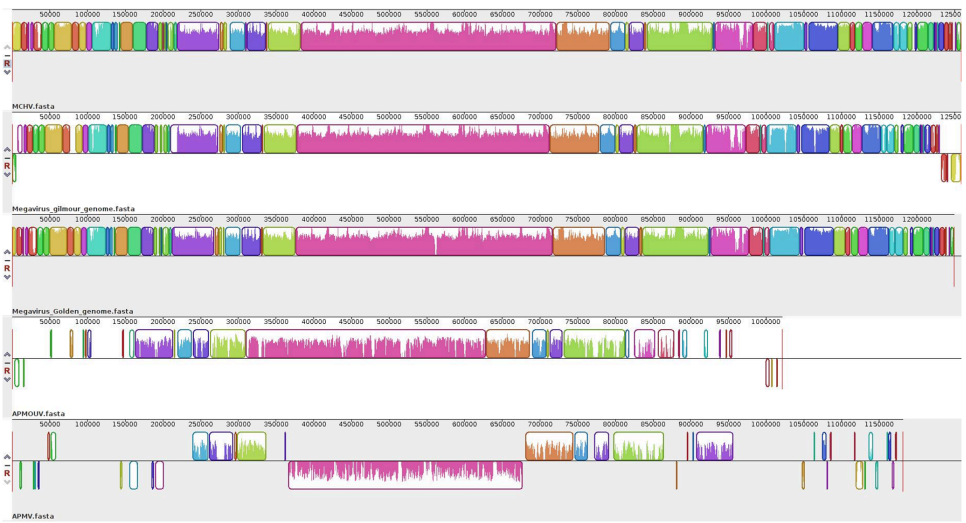


FIGURE 4 | Genome synteny analysis. Schematic genome alignment diagram obtained using the Mauve software package. The analysis was performed using the reference genome of APMV (NC_014649.1), Moumouvirus (NC_020104.1) and MCHV (NC_016072.1), besides the genome sequences of Br-mimiC isolates. The blocks illustrated above X axis are in positive strand (forward sense), while blocks below the X axis are in the negative strand (reverse sense).

predicted to encode only three tRNA molecules (Leu-TAA, Leu-CAA, and Trp-CCA). Taken together, these results confirm the isolation of the first mimiviruses of group C in Brazil.

In order to assess the gene encoding capacity of mimiviruses, we performed an updated pan-genome analysis (**Figure 3**) using all mimivirus genome data available in the NCBI genome database. This analysis will shows us the set of different proteins encoded by all mimiviruses, and will indicate whether the genetic complexity of this group has been fully addressed or not. For this analysis, only complete genome data sets were used, and the result showed a continuous increase in the pan-genome size reaching 2869 COGs, an improvement of 1740 new COGs compared with our previous analysis (Assis et al., 2015) that only considered genomes of mimiviruses of group A. Furthermore, breaks in this rising curve were observed for each new mimivirus representative of the lineages B and C; the number of COGs increased by 380

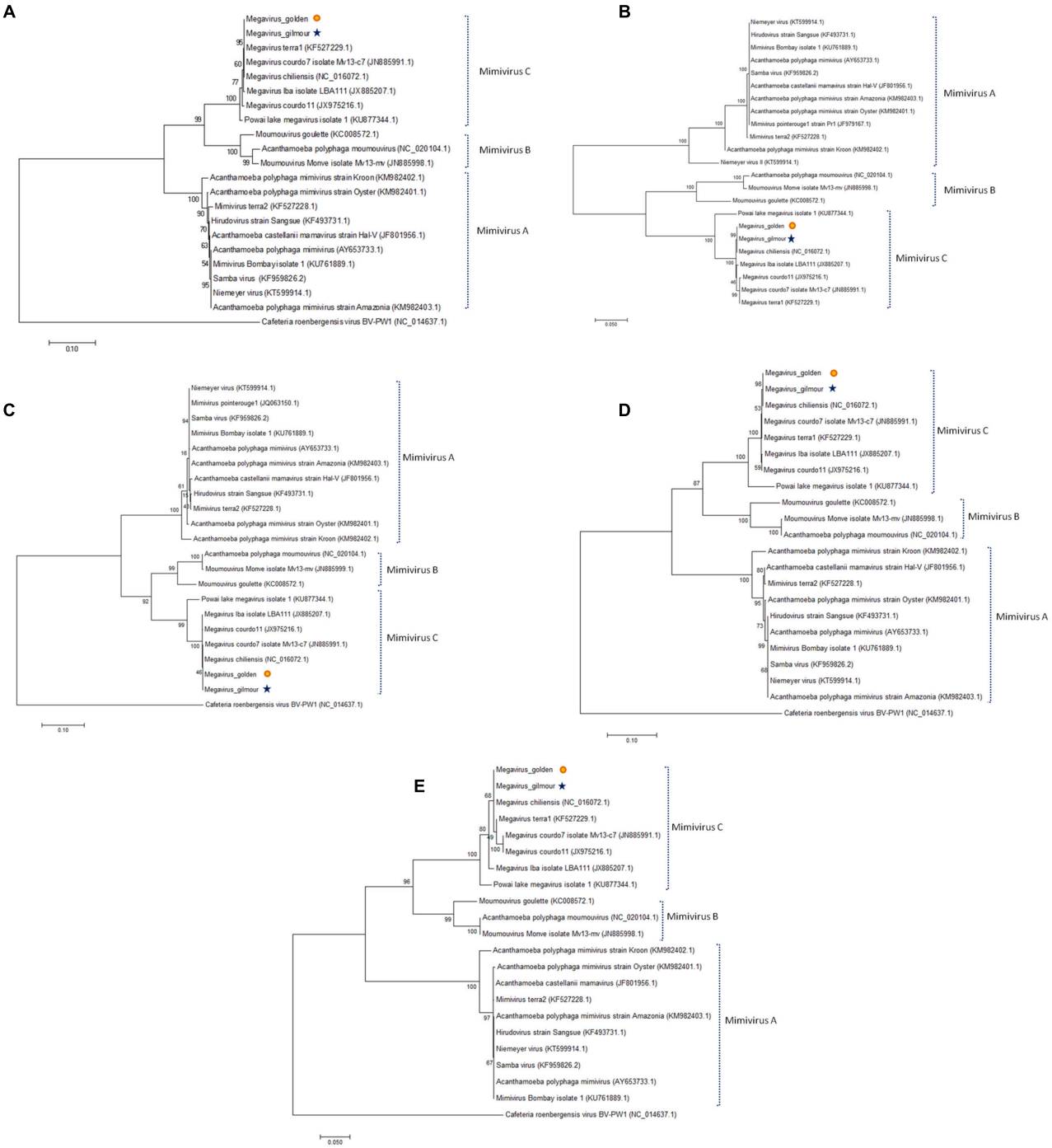


FIGURE 5 | Phylogenetic reconstruction of the NCLDV core genes for different mimiviruses isolates. Here we have done a phylogenetic analysis for many representatives of the different *Mimiviridae* lineages **(A–C)** including for the new Br-mimiC viruses, based on the nucleotide sequences of the five NCLDV core genes, namely the **(A)** major capsid protein, **(B)** the family B DNA polymerase, **(C)** the D6/D11 helicase, the **(D)** VV A18 helicase, and the **(E)** D5 primase-helicase.

from lineages A to B, and an additional increase of 208 COGs from lineages A and B to lineage C were observed. In addition, we observed a stabilization trend after the inclusion of the last four mimiviruses of group C, which included the Br-mimiC isolates.

Conversely, we observed a continuous decrease of the core genome after addition of new representatives. An abrupt reduction was only observed after inclusion of the first mimivirus of group B (268 COGs reduction), while a more discrete reduction was observed when sequences from mimiviruses C

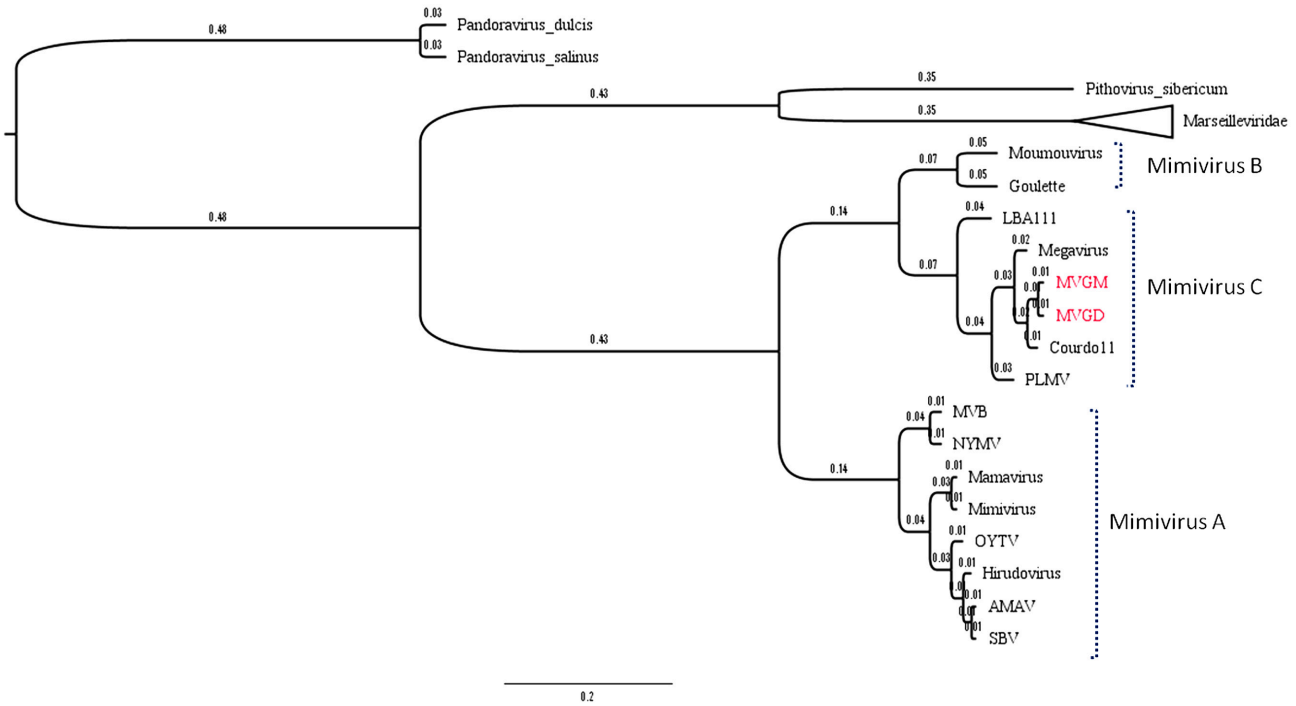


FIGURE 6 | Hierarchical clustering based on the pattern of gene presence/absence for the NCVOGs of different mimivirus isolates. Here we have done a hierarchical clustering tree, based on phyletic patterns, using a presence/absence matrix of 5,443 NCVOG from many representatives of the different *Mimiviridae* lineages (A, B, and C), including the new Br-mimiC viruses, highlighted in red. The distance from horizontal lines connecting the groups represents the time of divergence among the different mimiviruses isolates. The phylogenetic tree was generated using the MeV 4.8.1 tool.

were included (24 COGs reduction). Further, a stabilization trend of the core genome size was observed for the last five mimiviruses C, including Br-mimiC. In addition, we observed an intra-group divergence of 249 COGs among mimiviruses A, 487 COGs among mimiviruses B, and 563 COGs among mimiviruses C. Altogether, these results highlight a stabilization trend in the pan-genome and core genome evolution of amoeba-associated mimiviruses. In Addition our results showed a great divergence even among viruses from the same group (**Figure 3**).

Synteny Analysis

The synteny analysis showed very similar genome architectures for *Megavirus chilensis* (MCHV) and the Br-mimiC viruses (**Figure 4**). However, some divergences were observed among mimiviruses C, such as inversions and translocations at both extremities of the MVGM genome, while the MVGD genome better resembled the MCHV genome architecture than that of others. Furthermore, the genome of mimiviruses C showed a better co-linearity with less block brakes in their central region (from ∼250 to ∼950 kb) compared to what is observed at both extremities, which showed an increased number of shorter homologous regions. Curiously, the central region of mimiviruses C genomes showed an overall smaller similarity when compared to the extremities. In addition, these mimiviruses C presented a distinct genome macrosynteny from mimiviruses A and B.

Phylogeny

To better understand the evolutionary relationship between the Br-mimiC viruses and other mimiviruses, we performed phylogenetic analyses based on NCLDV core genes including the family B DNA polymerase, the VV A18 helicase, the D5 helicase, the D6/D11 helicase and the major capsid protein (**Figures 5A–E**). Furthermore, a hierarchical clustering tree (**Figure 6**), based on the phyletic patterns, was constructed using a presence-absence matrix of 5,443 NCVOG (clusters of orthologous genes shared by NCLDV). The phylogenetic trees recurrently clustered the Br-mimiC viruses into mimivirus group C, corroborating all previous analyses. The core genes-based trees showed the close relationship of Br-mimiC isolates to the MCHV isolate, the mimivirus of group C whose genome was first described, in 2011, and that was obtained from Chile. However, the phyletic tree, which highlights the gene presence/absence pattern, showed a close relationship of Br-mimiC with Courdo11 virus, isolated in 2010 by inoculating *Acanthamoeba* spp. with freshwater collected from a river of southeastern France.

DISCUSSION

In this work, we describe the isolation and genome features of the first two isolates of mimivirus group C from Brazil. Recently, we described the isolation of Samba virus, the first representative of family *Mimiviridae* in Brazil, belonging to

mimivirus group A (Campos et al., 2014). Subsequently, in Brazil, several mimiviruses and other giant viruses have been isolated in several biomes and a hospital respiratory-isolation facility, and mimivirus has been more recently detected in human sera (Dornas et al., 2014, 2015, 2017; Andrade et al., 2015; Boratto et al., 2015; dos Santos Silva et al., 2015). However, this is the first time that a mimivirus of group C is isolated in this country, which highlights the diversity of giant viruses in Brazil and how widespread these viruses are. Although the former member of mimivirus group C, *Mimivirus chilensis*, has been isolated in Chile, the remaining isolates of this group have frequently been isolated from environmental and clinical samples collected in Asia, Africa, and Europe (Pagnier et al., 2013). Thus, we believe that as new prospective studies are performed, new isolates might be discovered.

Even though they share many molecular features, as well as biological ones (data not shown), the Br-mimiC viruses can be recognized as two distinct isolates. The MVGM isolate has a genome ∼10 kb larger than the MVGD genome and encodes eight more ORFs than MVGD. The unique proteins of both Br-mimiC were mainly located at the extremities of both genomes, which have been described as suitable regions for horizontal gene transfers and duplication events in large and giant viruses, including in mimiviruses (Shackelton and Holmes, 2004; Colson et al., 2011a; Filee, 2015).

Even though the Br-mimiC viruses show ORFans in their genomes, which demonstrate the uniqueness of these isolates, there are notwithstanding many family ORFans present, which means that many genes are shared between mimiviruses but have no homolog in other organisms, and the majority of those genes remains functionally unresolved. Furthermore, we could see a still increasing pan-genome of the family *Mimiviridae* after the addition of Br-mimiC viral genome sequences, suggesting that new genes with unpredictable function are out there, yet to be discovered. In addition, the abrupt break in the trend of the core genome evolution after inclusion of lineage B sequences is in line with the fact that mimiviruses of lineages B and C are more related between each other than they are related to mimiviruses of lineage A, as also shown in the phylogeny reconstructions.

A more conserved synteny could be observed in the central region of all the mimiviruses genomes that were analyzed compared to the remaining part of the genomes. In contrast, the central region of mimiviruses C showed a lower mean similarity. The central region possesses the most ancient set of genes (Shackelton and Holmes, 2004), which have been subjected to long-term selective pressure during mimivirus evolutionary history. In contrast, termini regions of the genome more frequently incorporate new genes, and these recently acquired genes still have a more conserved profile. The phylogenetic analysis strongly corroborate all data presented above, indisputably showing the clustering of the new Br-mimiC isolates into mimivirus group C, closely related to *Megavirus chilensis*, the prototype of this group also isolated in South America. However, the phyletic analysis, which is based on gene presence/absence patterns that at least partially result from losses and gains, showed a better grouping of Br-mimiC viruses with the Courdo 11 virus isolate, which was isolated in 2010 from river water samples in France.

CONCLUSION

The discovery of the Br-mimiC viruses contributes to improving the understanding of mimiviral diversity and ubiquity. Nevertheless, the study of giant viruses is still at its beginning. Additional prospective studies must be conducted with the aim of discovering new relatives of these intriguing micro-organisms. Also, this study and others have showed a large number of sequences with unknown function, showing the need of studies focusing in the functional characterization of proteins encoded by the mimiviruses.

AUTHOR CONTRIBUTIONS

FA, PB, and AF-L: data collection and pan-genome analyses. FD, AF, RdS, and FC: samples collection and virus isolation. BLS, PC, and JA: study design. All authors wrote the paper and read its last version.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Andrade, K. R., Boratto, P. P., Rodrigues, F. P., Silva, L. C., Dornas, F. P., Pilotto, M. R., et al. (2015). Oysters as hot spots for mimivirus isolation. *Arch. Virol.* 160, 477–482. doi: 10.1007/s00705-014-2257-2

Arslan, D., Legendre, M., Seltzer, V., Abergel, C., and Claverie, J. M. (2011). Distant *Mimivirus* relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17486–17491. doi: 10.1073/pnas.1110889108

Assis, F. L., Bajrai, L., Abrahao, J. S., Kroon, E. G., Dornas, F. P., Andrade, K. R., et al. (2015). Pan-genome analysis of Brazilian lineage a amoebal Mimiviruses. *Viruses* 7, 3483–3499. doi: 10.3390/v7072782

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Bajrai, L. H., de Assis, F. L., Azhar, E. I., Jardot, P., Robert, C., Abrahao, J., et al. (2016). Saudi moumouvirus, the first group B *Mimivirus* isolated from Asia. *Front. Microbiol.* 7:2029. doi: 10.3389/fmicb.2016.02029

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29, 2607–2618. doi: 10.1093/nar/29.12.2607

Boratto, P. V., Arantes, T. S., Silva, L. C., Assis, F. L., Kroon, E. G., La Scola, B., et al. (2015). Niemeyer virus: a new *Mimivirus* group a isolate harboring a set of duplicated aminoacyl-tRNA synthetase genes. *Front. Microbiol.* 6:1256. doi: 10.3389/fmicb.2015.01256

Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D., and La Scola, B. (2013a). First isolation of a giant virus from wild *Hirudo medicinalis* leech: mimiviridae isolation in *Hirudo medicinalis*. *Viruses* 5, 2920–2930. doi: 10.3390/v5122920

Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D., and La Scola, B. (2013b). First isolation of a *Marseillevirus* in the diptera syrphidae *Eristalis tenax*. *Intervirology* 56, 386–394. doi: 10.1159/000354560

Boughalmi, M., Saadi, H., Pagnier, I., Colson, P., Fournous, G., Raoult, D., et al. (2013c). High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. *Environ. Microbiol.* 15, 2000–2007. doi: 10.1111/1462-2920.12068

Boyer, M., Madoui, M. A., Gimenez, G., La Scola, B., and Raoult, D. (2010). Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. *PLOS ONE* 5:e15530. doi: 10.1371/journal.pone.0015530

Campos, R. K., Boratto, P. V., Assis, F. L., Aguiar, E. R., Silva, L. C., Albarnaz, J. D., et al. (2014). Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. *Virol. J.* 11:95. doi: 10.1186/1743-422X-11-95

Chen, F., Mackey, A. J., Stoeckert, C. J. Jr., and Roos, D. S. (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 34, D363–D368. doi: 10.1093/nar/gkj123

Colson, P., de Lamballerie, X., Fournous, G., and Raoult, D. (2012). Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. *Intervirology* 55, 321–332. doi: 10.1159/000336562

Colson, P., Fournous, G., Diene, S. M., and Raoult, D. (2013). Codon usage, amino acid usage, transfer RNA and amino-acyl-tRNA synthetases in Mimiviruses. *Intervirology* 56, 364–375. doi: 10.1159/000354557

Colson, P., Gimenez, G., Boyer, M., Fournous, G., and Raoult, D. (2011a). The giant Cafeteria roenbergensis virus that infects a widespread marine phagocytic protist is a new member of the fourth domain of life. *PLOS ONE* 6:e18935. doi: 10.1371/journal.pone.0018935

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

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

Colson, P., Yutin, N., Shabalina, S. A., Robert, C., Fournous, G., La Scola, B., et al. (2011b). Viruses with more than 1,000 genes: mamavirus, a new *Acanthamoeba polyphaga* mimivirus strain, and reannotation of *Mimivirus* genes. *Genome Biol. Evol.* 3, 737–742. doi: 10.1093/gbe/evr048

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

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

Dornas, F. P., Boratto, P. V. M., Costa, G. B., Silva, L. C. F., Kroon, E. G., La Scola, B., et al. (2017). Detection of mimivirus genome and neutralizing antibodies in humans from Brazil. *Arch. Virol.* doi: 10.1007/s00705-017-3455-5 [Epub ahead of print].

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

Dornas, F. P., Rodrigues, F. P., Boratto, P. V., Silva, L. C., Ferreira, P. C., Bonjardim, C. A., et al. (2014). Mimivirus circulation among wild and domestic mammals, Amazon Region, Brazil. *Emerg. Infect. Dis.* 20, 469–472. doi: 10.3201/eid2003.131050

Dos Santos, R. N., Campos, F. S., Medeiros de Albuquerque, N. R., Finoketti, F., Correa, R. A., Cano-Ortiz, L., et al. (2016). A new marseillevirus isolated in Southern Brazil from *Limnoperna fortunei*. *Sci. Rep.* 6:35237. doi: 10.1038/srep35237

dos Santos Silva, L. K., Arantes, T. S., Andrade, K. R., Lima Rodrigues, R. A., Miranda Boratto, P. V., de Freitas Almeida, G. M., et al. (2015). High positivity of mimivirus in inanimate surfaces of a hospital respiratory-isolation facility, Brazil. *J. Clin. Virol.* 66, 62–65. doi: 10.1016/j.jcv.2015.03.008

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868. doi: 10.1073/pnas.95.25.14863

Filee, J. (2015). Genomic comparison of closely related Giant Viruses supports an accordion-like model of evolution. *Front. Microbiol.* 6:593. doi: 10.3389/fmicb.2015.00593

Fischer, M. G., Allen, M. J., Wilson, W. H., and Suttle, C. A. (2010). Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19508–19513. doi: 10.1073/pnas.1007615107

International Committee on Taxonomy of Viruses [ICTV] (2017). Available at: https://talk.ictvonline.org/ [accessed September, 2017].

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

La Scola, B., Campocasso, A., N'Dong, R., Fournous, G., Barrassi, L., Flaudrops, C., et al. (2010). Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. *Intervirology* 53, 344–353. doi: 10.1159/000312919

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* 455, 100–104. doi: 10.1038/nature07218

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

Legendre, M., Arslan, D., Abergel, C., and Claverie, J. M. (2012). Genomics of Megavirus and the elusive fourth domain of life. *Commun. Integr. Biol.* 5, 102–106. doi: 10.4161/cib.18624

Levasseur, A., Bekliz, M., Chabriere, E., Pontarotti, P., La Scola, B., and Raoult, D. (2016). MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature* 531, 249–252. doi: 10.1038/nature17146

Mutsafi, Y., Zauberman, N., Sabanay, I., and Minsky, A. (2010). Vaccinia-like cytoplasmic replication of the giant *Mimivirus*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5978–5982. doi: 10.1073/pnas.0912737107

Pagnier, I., Reteno, D. G., Saadi, H., Boughalmi, M., Gaia, M., Slimani, M., et al. (2013). A decade of improvements in Mimiviridae and Marseilleviridae isolation from amoeba. *Intervirology* 56, 354–363. doi: 10.1159/000354556

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004). The 1.2-megabase genome sequence of *Mimivirus*. *Science* 306, 1344–1350. doi: 10.1126/science.1101485

Rodriguez-R, L. M., and Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe* 9, 111–118. doi: 10.1128/microbe.9.111.1

Saadi, H., Pagnier, I., Colson, P., Cherif, J. K., Beji, M., Boughalmi, M., et al. (2013a). First isolation of *Mimivirus* in a patient with pneumonia. *Clin. Infect. Dis.* 57, e127–e134. doi: 10.1093/cid/cit354

Saadi, H., Reteno, D. G., Colson, P., Aherfi, S., Minodier, P., Pagnier, I., et al. (2013b). Shan virus: a new *Mimivirus* isolated from the stool of a Tunisian patient with pneumonia. *Intervirology* 56, 424–429. doi: 10.1159/000354564

Sambrook, J., and Russell, R. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd Edn. Cold Spring Harbor, NY: Cold spring harbor laboratory press.

Schattner, P., Brooks, A. N., and Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33, W686–W689. doi: 10.1093/nar/gki366

Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., et al. (2017). Giant viruses with an expanded complement of translation system components. *Science* 356, 82–85. doi: 10.1126/science.aal4657

Shackelton, L. A., and Holmes, E. C. (2004). The evolution of large DNA viruses: combining genomic information of viruses and their hosts. *Trends Microbiol.* 12, 458–465. doi: 10.1016/j.tim.2004.08.005

Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L., and Raoult, D. (2007). Ultrastructural characterization of the giant volcano-like virus factory of *Acanthamoeba polyphaga Mimivirus*. *PLOS ONE* 2:e328. doi: 10.1371/journal.pone.0000328

Takemura, M., Mikami, T., and Murono, S. (2016). Nearly complete genome sequences of two *Mimivirus* strains isolated from a Japanese freshwater pond and river mouth. *Genome Announc.* 4:e01378-16. doi: 10.1128/genomeA.01378-16

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Yoosuf, N., Pagnier, I., Fournous, G., Robert, C., La Scola, B., Raoult, D., et al. (2014a). Complete genome sequence of Courdo11 virus, a member of the family Mimiviridae. *Virus Genes* 48, 218–223. doi: 10.1007/s11262-013-1016-x

Yoosuf, N., Pagnier, I., Fournous, G., Robert, C., Raoult, D., La Scola, B., et al. (2014b). Draft genome sequences of Terra1 and Terra2 viruses, new members

of the family Mimiviridae isolated from soil. *Virology* 45, 125–132. doi: 10.1016/j.virol.2013.12.032

Yoosuf, N., Yutin, N., Colson, P., Shabalina, S. A., Pagnier, I., Robert, C., et al. (2012). Related giant viruses in distant locations and different habitats: *Acanthamoeba polyphaga* moumouvirus represents a third lineage of the Mimiviridae that is close to the megavirus lineage. *Genome Biol. Evol.* 4, 1324–1330. doi: 10.1093/gbe/evs109

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

Yutin, N., Wolf, Y. I., Raoult, D., and Koonin, E. V. (2009). Eukaryotic large nucleocytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol. J.* 6, 01–13. doi: 10.1186/1743-422X-6-223

Zauberman, N., Mutsafi, Y., Halevy, D. B., Shimoni, E., Klein, E., Xiao, C., et al. (2008). Distinct DNA exit and packaging portals in the virus *Acanthamoeba polyphaga* mimivirus. *PLOS Biol.* 6:e114. doi: 10.1371/journal.pbio.0060114

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

frontiers in Microbiology

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

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

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

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

OPEN ACCESS

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

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

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

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

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

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

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

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

INTRODUCTION

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

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

MATERIALS AND METHODS

Virus Isolation, Production, and Purification

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

Genome Sequencing

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

Genome Assembly

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

Transcriptome Sequencing of *Pandoravirus massiliensis*

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

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

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

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

Proteome Analysis of *Pandoravirus massiliensis*

Preparation of the Total Proteins of the Virus

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

Two-Dimensional Gels

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

Mass Spectrometry

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

Genome Annotation

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

Transcriptomic Analysis for *Pandoravirus massiliensis*

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

Search for Transposable Elements

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

Phylogenetic Analyses and Hierarchical Clustering

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

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

RESULTS

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

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

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

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

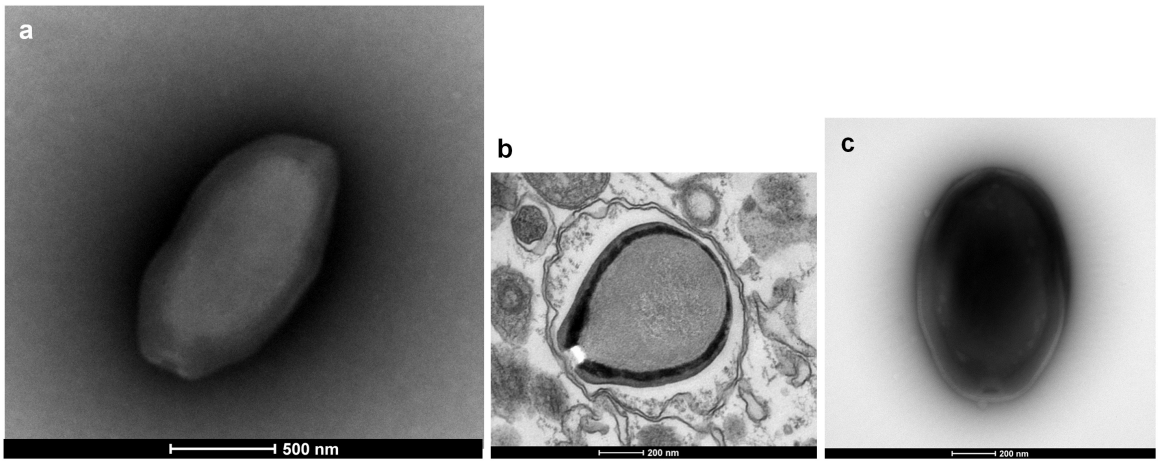


FIGURE 1 | Electron microscopy pictures of pandoravirus isolates by negative staining **(a,c)** or after inclusion **(b)**. **(a)** *Pandoravirus massiliensis*; **(b)** *Pandoravirus pampulha*; **(c)** *Pandoravirus braziliensis*.

TABLE 1 | Main features of the six pandoravirus genomes.

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

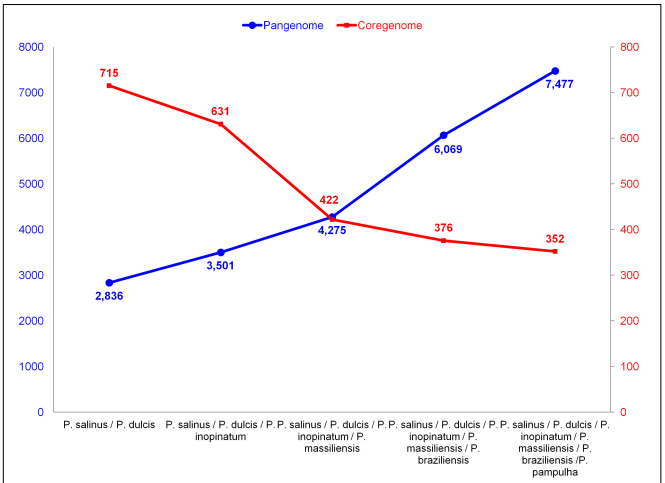


FIGURE 2 | Evolution of the size of the pangenome and core genome of pandoraviruses with the description of each of the six pandoravirus isolates.

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

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

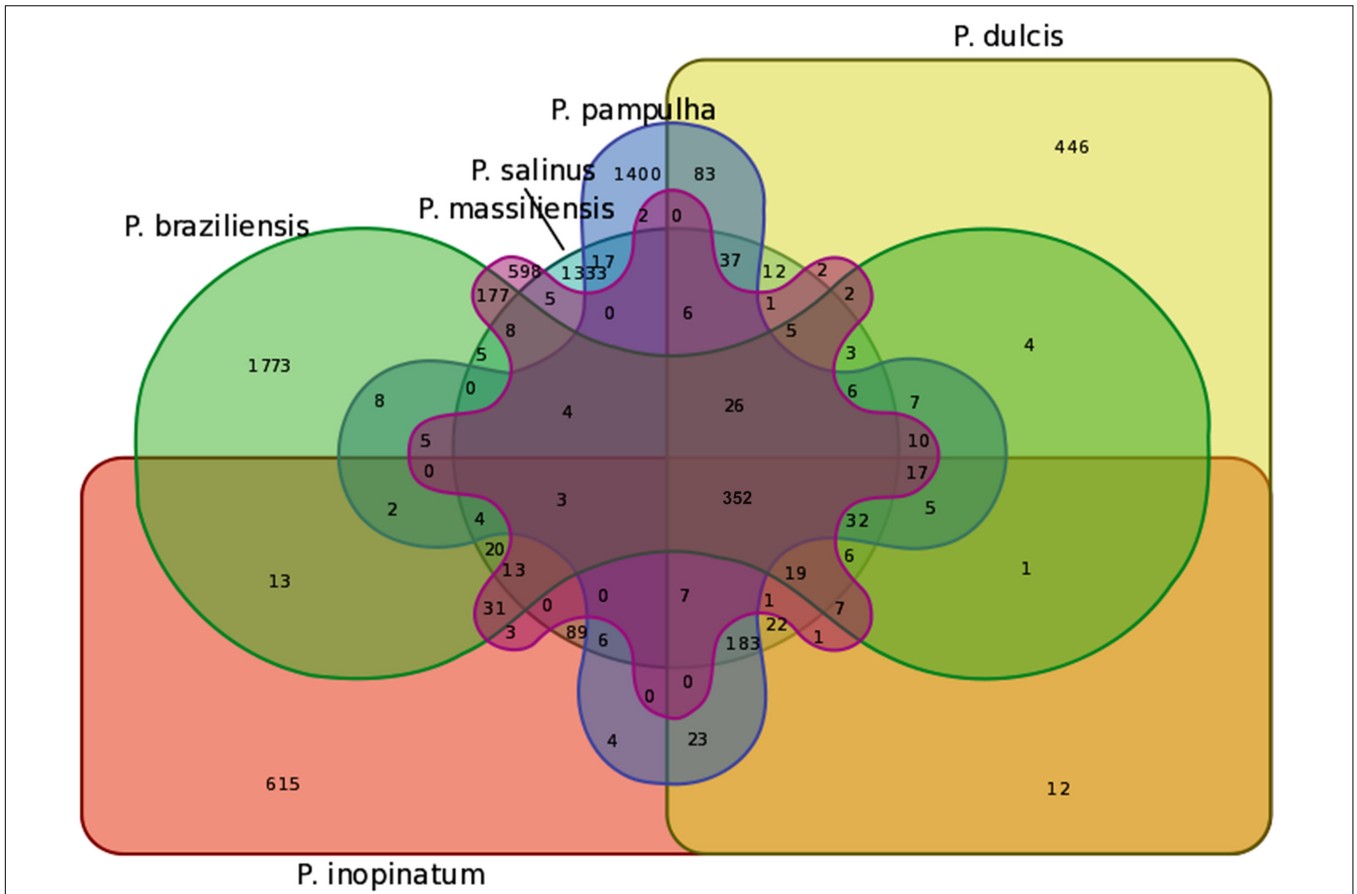


FIGURE 3 | Venn diagram of genes shared and not shared between the gene contents of the six pandoravirus isolates. Venn diagram was built using the following online tool: http://bioinformatics.psb.ugent.be/webtools/Venn/.

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

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

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

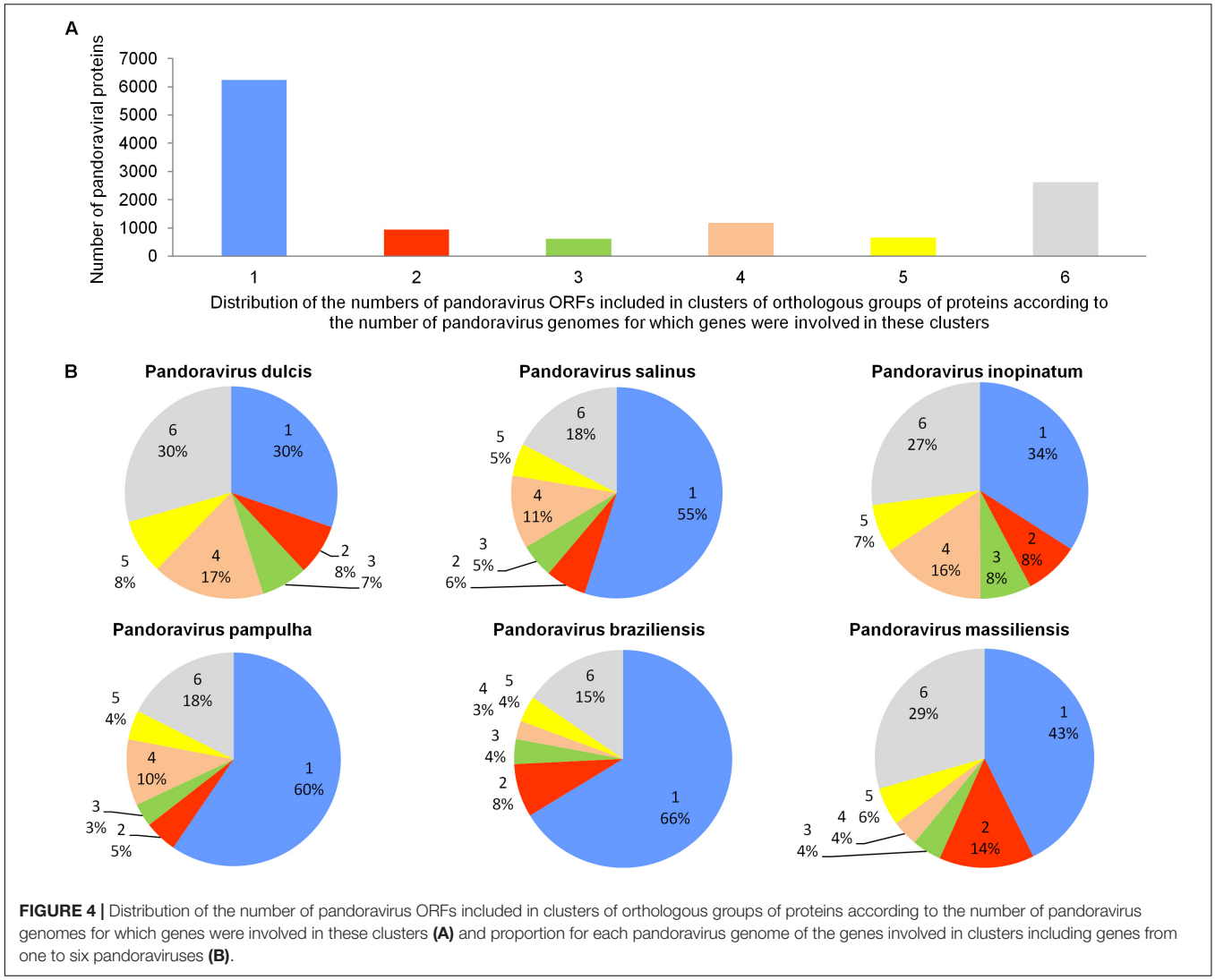


FIGURE 4 | Distribution of the number of pandoravirus ORFs included in clusters of orthologous groups of proteins according to the number of pandoravirus genomes for which genes were involved in these clusters **(A)** and proportion for each pandoravirus genome of the genes involved in clusters including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

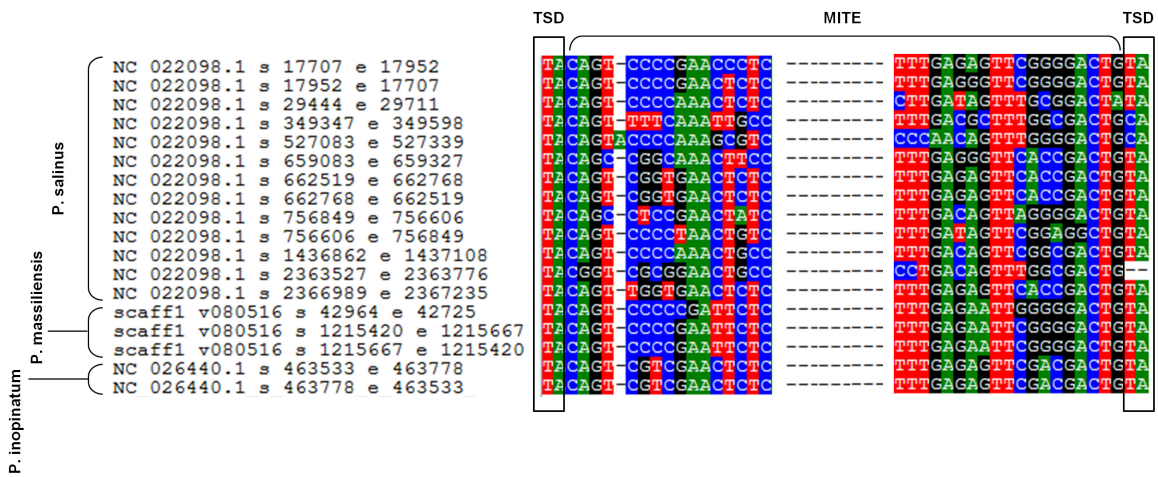


FIGURE 5 | Alignment by muscle of the 13 full-length miniature inverted repeat transposable elements (MITEs) described in the genome of *Pandoravirus salinus* (Sun et al., 2015) and those detected in the genomes of *P. massiliensis* and *P. inopinatum*. Alignment is limited to regions including the flanking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

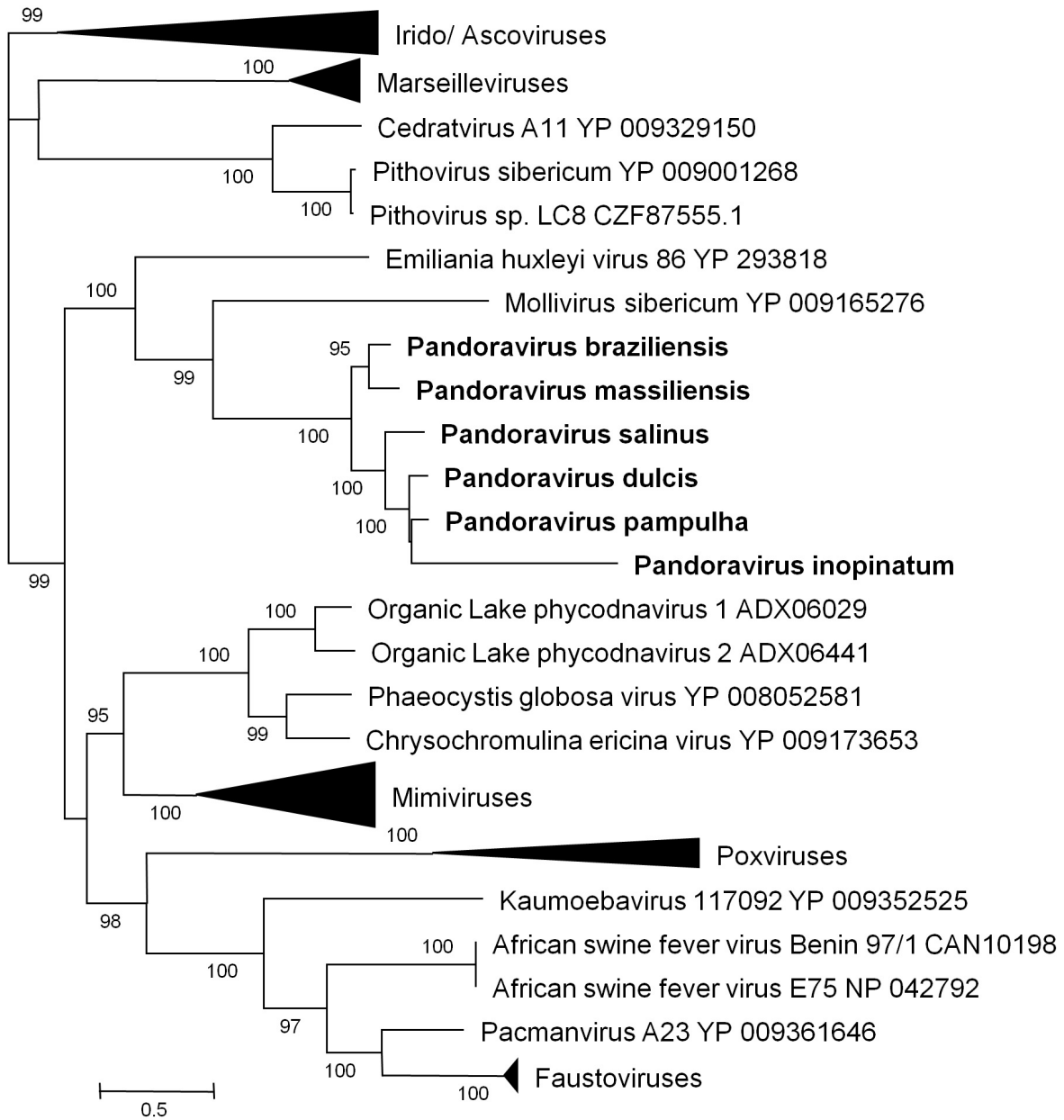


FIGURE 6 | Phylogenetic reconstruction based on amino acid sequences of the DNA-dependent RNA polymerase subunit 1 from representatives of megavirales. Phylogenetic tree was drawn using the maximum likelihood model with the FastTree program (Price et al., 2010).

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

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

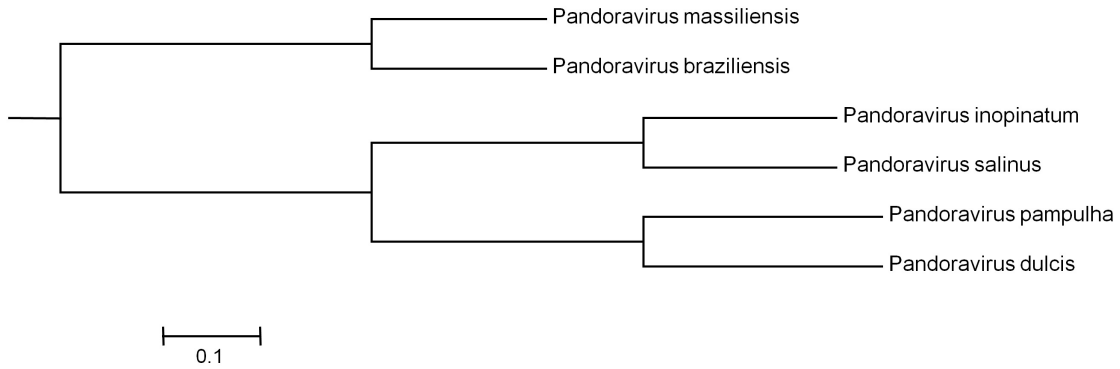


FIGURE 7 | Hierarchical clustering based on the presence/absence patterns of clusters of orthologous groups of proteins of megavirales members in the genomes of pandoraviruses.

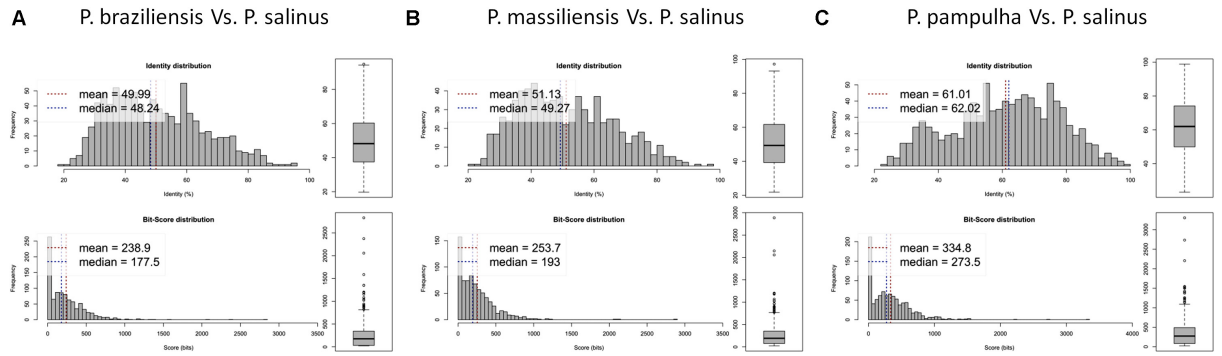


FIGURE 8 | Average amino acid identity between ORFs predicted for pandoravirus genomes. For each comparison, estimates were obtained using both best hits and reciprocal best hits between two sets of proteins from a pandoravirus newly described here and *Pandoravirus salinus*. **(A)** *Pandoravirus braziliensis*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

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

DISCUSSION

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

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

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

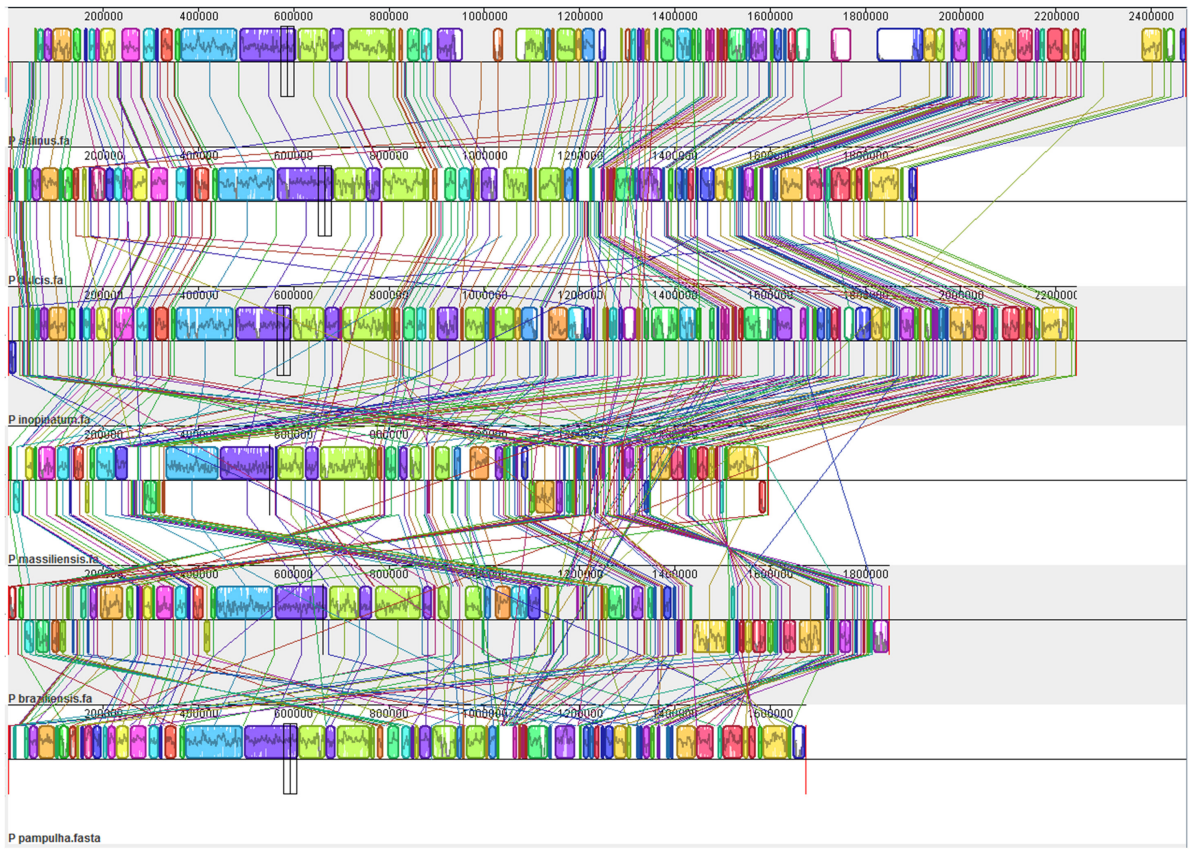


FIGURE 9 | Whole genome alignment of pandoravirus genomes by the MAUVE program (Darling et al., 2010).

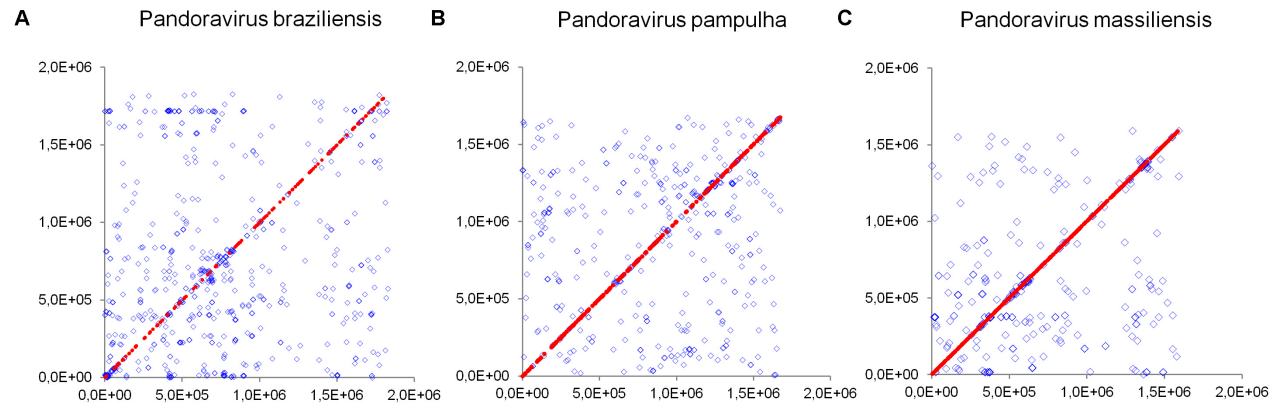


FIGURE 10 | Distribution of core genes and paralogous genes along the pandoravirus genomes. **(A)** *Pandoravirus braziliensis*; **(B)** *Pandoravirus pampulha*; **(C)** *Pandoravirus massiliensis*. Core genes are indicated by red dots; paralogous genes are indicated by blue diamonds.

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

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

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

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

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

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by the French Government under the "Investissements d'avenir" (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research) (reference: Méditerranée Infection 10-IAHU-03), and by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

ACKNOWLEDGMENTS

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

SUPPLEMENTARY MATERIAL

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

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

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

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (2015). In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting *Acanthamoeba*. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5327–E5335. doi: 10.1073/pnas.1510795112

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

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

Okamoto, K., Miyazaki, N., Song, C., Maia, F. R. N. C., Reddy, H. K. N., Abergel, C., et al. (2017). Structural variability and complexity of the giant *Pithovirus sibericum* particle revealed by high-voltage electron cryo-tomography and energy-filtered electron cryo-microscopy. *Sci. Rep.* 7, 13291–13390. doi: 10.1038/s41598-017-13390-4

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 23 October 2018
doi: 10.3389/fmicb.2018.02534

Faustovirus E12 Transcriptome Analysis Reveals Complex Splicing in Capsid Gene

*Amina Cherif Louazani, Emeline Baptiste, Anthony Levasseur, Philippe Colson and Bernard La Scola**

Assistance Publique – Hôpitaux de Marseille (AP-HM), Microbes, Evolution, Phylogeny and Infection (MEΦI), Institut Hospitalo-Universitaire (IHU) Méditerranée Infection, Institut de Recherche pour le Développement IRD 198, Aix-Marseille Université UM63, Marseille, France

Faustoviruses are the first giant viruses of amoebae isolated on *Vermamoeba vermiformis*. They are distantly related to African swine fever virus, the causative agent of lethal hemorrhagic fever in domestic pigs. Structural studies have shown the presence of a double protein layer encapsidating the double-stranded DNA genome of Faustovirus E12, the prototype strain. The major capsid protein (MCP) forming the external layer has been shown to be 645-amino acid-long. Unexpectedly, its encoding sequence has been found to be scattered along a 17 kbp-large genomic region. Using RNA-seq, we studied expression of Faustovirus E12 genes at nine time points over its entire replicative cycle. Paired-end 250 bp-long read sequencing on MiSeq instrument and double-round spliced alignment enabled the identification of 26 different splice-junctions. Reads corresponding to junctions represented 2% of mapped reads and mostly matched with the predicted MCP encoding sequences. Moreover, our study enabled describing a 1,939 bp-long transcript that corresponds to the MCP, delineating 13 exons. At least two types of introns coexist in the MCP gene: group I introns that can self-splice (n = 5) and spliceosome-like introns with non-canonical splice sites (n = 7). All splice-sites were non-canonical with five types of donor/acceptor splice-sites among which AA/TG was the most frequent association.

Keywords: giant virus, faustovirus, transcriptome, capsid, splicing

OPEN ACCESS

Edited by:
Erna Geessien Kroon, Universidade Federal de Minas Gerais (UFMG), Brazil

Reviewed by:
Juliana Cortines, Universidade Federal do Rio de Janeiro, Brazil
Masaharu Takemura, Tokyo University of Science, Japan

***Correspondence:**
Bernard La Scola
bernard.la-scola@univ-amu.fr

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

Received: *30 July 2018*
Accepted: *04 October 2018*
Published: *23 October 2018*

Citation:
Cherif Louazani A, Baptiste E, Levasseur A, Colson P and La Scola B (2018) Faustovirus E12 Transcriptome Analysis Reveals Complex Splicing in Capsid Gene. Front. Microbiol. 9:2534. doi: 10.3389/fmicb.2018.02534

INTRODUCTION

Faustoviruses are the first giant viruses of amoebae isolated using *Vermamoeba vermiformis* as cellular culture support (Reteno et al., 2015). Their capsids are icosahedral and virions are 200–240 nm large (Benamar et al., 2016). These viruses are distantly related to African swine fever virus, the causative agent of lethal hemorrhagic fever in domestic pigs (Alonso et al., 2018) and single species of family *Asfarviridae* (Iyer et al., 2001). In addition, two other faustovirus relatives have recently been described. Kaumoebavirus, also isolated on *V. vermiformis*, stands phylogenetically outside the asfarvirus–faustovirus group (Bajrai et al., 2016). Pacmanvirus, isolated on *Acanthamoeba castellanii,* is nested in phylogenetic analyses between asfarviruses and

faustoviruses (Andreani et al., 2017). So far, 11 faustovirus isolates have been isolated, in all cases from sewage samples collected in France, Lebanon and Senegal (Cherif Louazani et al., 2017). Faustovirus-like sequences were also identified in metagenomes generated from arthropods as well as from febrile patients, healthy people, and from rodents (Temmam et al., 2015).

To better characterize the genomic diversity of faustoviruses, the genomes of the 11 isolates have been sequenced and annotated. These double-stranded DNA genomes contain between 456 and 491 kilobase pairs (kbp), have a G + C content comprised between 36.2 and 39.6%, and were predicted to encode between 457 and 519 genes (Benamar et al., 2016). Four lineages could be inferred from phylogenetic analyses of the core genome, with no clustering of the strains according to their geographical origin (Benamar et al., 2016; Cherif Louazani et al., 2017). For all these isolates, many hypothetical proteins were predicted, for which no function could be inferred due to the absence of recognizable homologs or conserved domains, their number being 148 among proteins encoded by the core genes.

In Faustovirus E12, the prototype virus of this group, proteomic analyses confirmed the presence in mature virions of 162 (33%) of the predicted proteins (Reteno et al., 2015). Moreover, cryo-electron microscopy showed the presence of a double protein layer encapsidating its genome (Klose et al., 2016). The major capsid protein (MCP) forming its external protein layer has been shown to be 645-amino acid-long. In addition, it folds into the double jelly roll motif that is characteristic of the capsid proteins of large nucleo-cytoplasmic double-stranded DNA viruses (NCLDV), a group of viral families that comprises the *Asfarviridae* family (Iyer et al., 2001). Strikingly, the sequences encoding the Faustovirus E12 MCP appeared to be scattered along a 17 kbp-large genomic region, with fragments located in both annotated and unannotated ORFs. This observation suggested that Faustovirus E12 uses an extended splicing during the expression of its MCP (Reteno et al., 2015; Klose et al., 2016).

In silico gene finding approaches have limitations in identifying genes, especially those that undergo post-transcriptional modifications or are present in the genomes of non-model organisms (Klasberg et al., 2016). The RNA-seq technology is particularly helpful in such cases. Using high throughput sequencing, RNA-seq allows high resolution identification of whole genome transcripts, of splicing events and splice junctions. It delineates the transcriptional structure of genes, and provides interesting information on gene expression levels and kinetics (Wang et al., 2009). Thus, previous studies of giant virus transcriptomes used RNA-seq to validate gene predictions and determine the precise 5′ and 3′ UTR structures of transcripts (Legendre et al., 2014, 2015). For Mimivirus, this approach increased the gene repertoire of 49 genes and detected a new component of the transcription apparatus (Legendre et al., 2011).

In the present study, we provide a comprehensive view of Faustovirus E12 genes expression through massive parallel sequencing of the total RNA-derived cDNA. We put a special focus on the identification of splicing events in the transcription process of the MCP encoding gene over the entire replicative cycle.

MATERIALS AND METHODS

A flowchart summarizing the main steps used in this study is presented in **Figure 1**.

Data Acquisition

Virus Production and Infection Cycle

Faustovirus E12 was produced on *V. vermiformis* (strain CDC19) as in a previously described procedure (Reteno et al., 2015). Briefly, confluent monolayers of amoebae in Peptone-Yeast extract-Glucose (PYG) medium incubated at 28°C were rinsed with Page's Amoeba Saline buffer (PAS) and centrifugated twice at 720 × *g* for 10 min, then put in a starvation medium at an adjusted concentration of 10^6 cells/mL. The amoebae were then incubated at 30°C with a viral suspension at a MOI of five until complete cell lysis. The culture supernatant was then filtered at 0.45 μm to eliminate cellular debris and the filtrate was titrated by limited dilution assay.

For the interrupted infection cycle, adherent *V. vermiformis* incubated in PYG medium were put in contact with viral suspension at a MOI of 100. After incubation at 30°C for 1 h, the supernatant was removed, and the cultures were gently rinsed three times with PAS to eliminate excess virus. This marked time 0 (T0). For later time points, infected and rinsed amoebae were incubated at 30°C in PYG. Infected cells were pelleted by centrifugation at 720 × *g* for 10 min and were stored at −80°C in PBS.

In total, we realized two infection cycles with the following post-infection time points in duplicate: (*t* = 0, 15 min, 90 min, 3 h, 6 h, and 8 h), hereafter referred to as T0min-1, T15min-1, T90min-1, T3H-1, T6H-1, T8H-1 for cycle 1 and T0min-2, T15min-2, T90min-2, T3H-2, T6H-2, and T8H-2 for cycle 2. The second cycle included three additional late time points (*t* = 11 h, 17 h, and 20 h): T11H, T17H, and T20H.

RNA Extraction and cDNA Sequencing

RNA was extracted using the RNeasy mini kit (Cat No: 74104, Qiagen, France) according to the manufacturer's instructions. Total RNA was eluted in a 50 μL volume of RNase-free water. RNaseOUT (Thermo Fisher Scientific, France) was added to the elute to prevent RNA degradation. Genomic DNA contamination was checked using a PCR system targeting Faustovirus E12 DNA (forward primer: TCGGCATCAATCGCCTTATAG; reverse primer: GGCCAGAAGGGTCATTAACA). Two cycles of 30 min-DNase treatment using TURBO DNase (Invitrogen, France) incubation at 37°C were performed on the samples to achieve absence of DNA contamination. RNeasy MinElute Cleanup Kit (Qiagen) was used to purify DNA-free total RNA, using the manufacturer's protocol with an RNA elution volume of 14 μL in RNase-free water.

The extracted total RNAs were reverse transcribed into cDNA using random primers with the SuperScript VILO Synthesis Kit (Invitrogen, France). cDNA amplicons were purified with

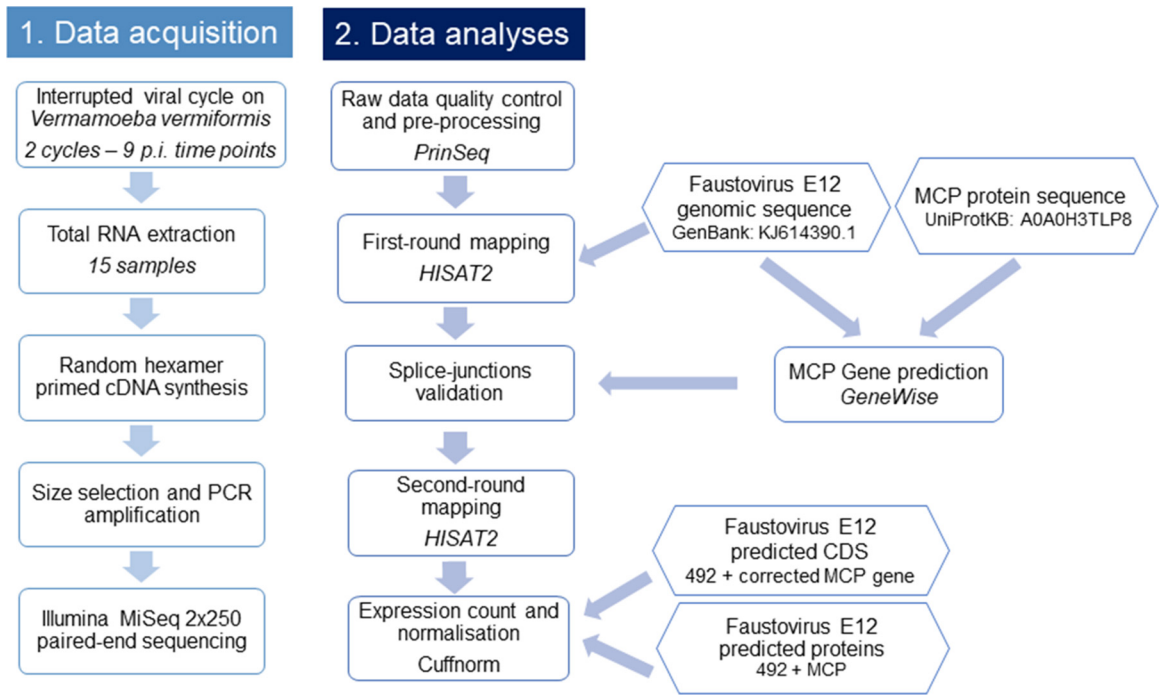


FIGURE 1 | Flowchart illustrating the workflow of this study. This flowchart shows the general pipeline of this RNA-seq study, starting from sample preparation and RNA extraction to cDNA sequencing and data analyses. The biological interpretation of expression count was possible through the functional categories clustering of expressed genes.

the Agencourt AMPure XP system (Beckman Coulter Inc., CA, United States). Two sets of purified cDNA corresponding to the early and a complete Faustovirus E12 infection cycle were sequenced on a MiSeq instrument with the 2-bp × 250-bp paired-end strategy, using Nextera XT DNA sample prep kit (Illumina Inc., CA, United States). Quantified cDNAs were fragmented, tagged, then barcoded through limited cycle PCR amplification (12 cycles). After purification on Agencourt AMPure XP beads (Beckman Coulter Inc., CA, United States), the libraries were normalized on specific beads and pooled for sequencing. Each set was loaded on a separate flowcell.

Data Analyses

Quality Control and Pre-processing of Reads

The raw data of paired-end reads were adapter trimmed. Adapter-free reads were checked for quality using PrinSeq web-version 0.20.1 (Schmieder and Edwards, 2011). Reads with over 10% Ns were filtered out. PolyA/T tails of over seven nucleotides (nt) were trimmed. Reads were quality trimmed from 5′-end with a sliding window of four and a step of three, with a mean Phred-scaled quality score cutoff of 20.

Study of Faustovirus Genes Expression

To identify potential splicing events in Faustovirus E12, we used a two-round alignment approach with a spliced-mapper: first, both pre-processed paired reads and singleton were mapped against the genomic sequence of Faustovirus E12 (GenBank accession no. KJ614390.1) using HISAT2 with minimum and maximum size of introns set to 20 and 5,000 bp (Kim et al., 2015). Spliced reads were extracted, and junctions manually validated using the Gene BED To Exon/Intron/Codon BED expander (Galaxy Version 1.0.0) (Afgan et al., 2016) and the Integrative Genomics Viewer (IGV) tool (Thorvaldsdottir et al., 2013). Junctions supported by at least two reads were included as known junctions in the second alignment round.

For each time point, reads mapping to the viral genes were quantified and the counts normalized with the geometric method using Cuffnorm (Galaxy Version 2.2.1.1) (Trapnell et al., 2010). For t = 0 to t = 8 h p.i., for which two biological replicates were available, both replicates were used as entries for a common normalized count (T0-c to T8H-c). To study the functional profile of the genes expressed during the replicative cycle, a BLASTp (Altschul et al., 1997) search of Faustovirus E12 annotated ORFs was performed against the Nucleo-Cytoplasmic Virus Orthologous Groups (NCVOGs) proteins database[1] (Yutin et al., 2014). Hits with e-values below 1e-03 were considered significant and assigned to their corresponding NCVOGs. A weighted average of expressed genes in Fragments Per Kilobase of transcript per Million mapped reads (FPKM) was calculated for each functional category at each time point.

Proteins of the African swine fever virus (ASFV) identified in the purified particles (Alejo et al., 2018) were searched for homologs in Faustovirus E12 using BLASTp (Altschul et al., 1997) with 1e-03 as cutoff.

[1] ftp://ftp.ncbi.nih.gov/pub/wolf/COGs/NCVOG/

Study of the Major Capsid Protein Encoding Gene in Faustovirus E12

The 645-amino-acid protein sequence of the Faustovirus E12 MCP (UniProtKB accession no.: A0A0H3TLP8) was used to predict coding regions in the viral genome, using GeneWise (online version: wise2-4-1) (Li et al., 2015) with the GeneWise 623 algorithm, the flat null model and modeled splice sites as entry parameters. Predicted positions of exons were manually curated using information from junction reads. The coordinates of the corresponding junctions have been added to the file of known splice junctions for the second-round alignment of total RNA-derived cDNA.

RESULTS

Faustovirus E12 Gene Expression

The transcriptome sequencing of Faustovirus E12-infected *V. vermiformis* resulted in 8,909,144 read pairs distributed over nine time points with two biological replicates corresponding to t = 0 min, 15 min, 90 min, 3 h, 6 h, and 8 h, and one replicate for t = 11, 17, and 20 h. After quality control, pre-processing and double-round mapping, reads corresponding to Faustovirus E12 represented <1% of the total number of generated reads, yet covering 93.5% of the genome positions with at least one read in at least one dataset. Single-base-resolution coverage maps across the genome for datasets of both cycles are reported in **Figure 2** and **Supplementary Figure 1**. We observed a gradual increase in genome coverage during the replication cycle, illustrating an active transcription process starting early after infection. Two major shifts in coverage peaks profiles were observed after t = 90 min and t = 8 h, marking transitions from early to intermediate and from intermediate to late infection time points.

We detected that Faustovirus E12 expresses during its replicative cycle 90% (445/492) of its predicted genes including all but two genes that were assigned an NCVOG ID (116/118) (**Supplementary Table 1**). These two genes are a putative metal-dependent hydrolase (PRJ_Fausto_00294) and an uncharacterized protein (PRJ_Fausto_00234). Genes related to DNA replication, recombination and repair; nucleotide metabolism and transcription and RNA processing were expressed early and throughout the whole cycle. These include a hydrolase and a putative P-loop containing nucleoside triphosphate hydrolase, the ribonucleotide reductase small and large subunits and the hypothetical protein (PRJ_Fausto_00128) containing a Rho factor transcription termination domain.

A large amount (32–75%) of the transcripts detected at early time points and up to 6 h p.i. corresponded to uncharacterized or poorly characterized proteins. Among the early expressed genes, we also found genes predicted to be involved in the ubiquitin-proteasome pathway and in host response regulation notably ankyrin repeats and membrane occupation and recognition nexus (MORN) repeat containing proteins. DNA directed RNA polymerase subunits are expressed starting from 90 min p.i. along with the transcription factor S-II (TFIIS), the mRNA capping enzyme and the translation initiation factor SUI1. The first transcripts corresponding to the MCP appeared at 3 h p.i while genes related to virion structure and morphogenesis were expressed starting from 6 h p.i. with increasing abundance in the late times. From 8 h p.i., the majority (50–73%) of the transcripts corresponded to proteins involved in virion structure and morphogenesis (**Figure 3**). **Table 1** lists Faustovirus E12 genes predicted to encode for homolog proteins to those detected

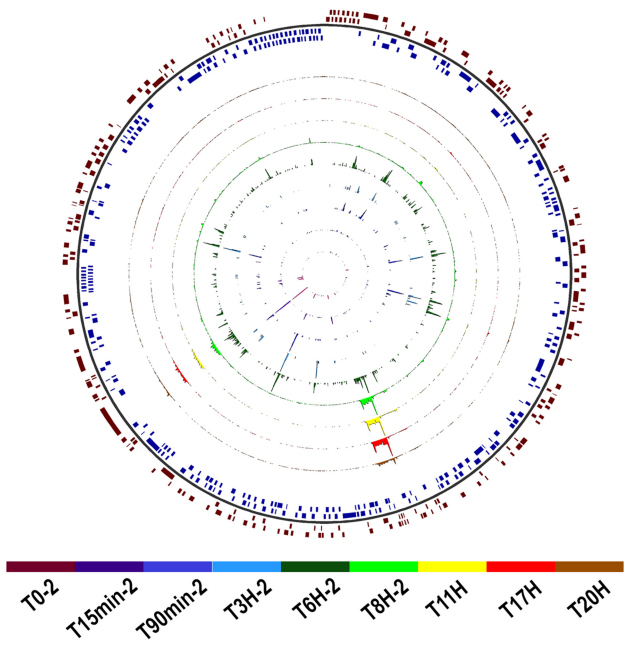


FIGURE 2 | Map of Faustovirus E12 genome coverage during the replication cycle. The predicted protein coding sequences are represented on the external circle in *red* and *blue boxes* for the forward and reverse strand, respectively. The single base resolution coverage for each time point is reported in the colored concentric circles from t = 0 to 20 h for the complete replication cycle samples set. Position 0 is at the 12 o'clock position.

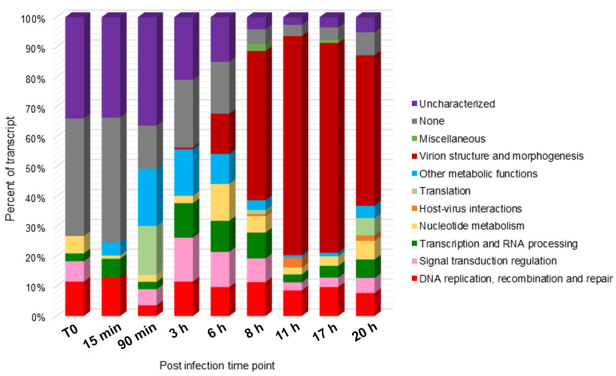


FIGURE 3 | Functional categories distribution of expressed genes during the replication cycle of Faustovirus E12. Faustovirus E12 ORFs were assigned a functional category based on sequence homology with the Nucleo-Cytoplasmic Virus Orthologous Groups (NCVOG) proteins database. The ratio of expression per functional category is reported for each time point.

TABLE 1 | ASFV virion-forming proteins with homologs detected in the transcriptome of Faustovirus E12.

ASFV gene	ASFV protein	Function	Faustovirus E12 homolog protein	T8H	T11H	T17H	T20H
B646L	Major capsid protein p72	Morphogenesis	MCP	+	+	+	+
CP2475L	Polyprotein pp220	Morphogenesis	AIB52024	+	+	+	+
CP530R	Polyprotein pp62	Morphogenesis	AIB52025	+	+	+	+
S273R	Polyprotein processing protease	Morphogenesis	AIB52094	+	+	+	+
O174L	DNA polymerase X	DNA integrity	AIB52077	+			
E296R	AP endonuclease	DNA integrity	AIB52085	+	+	+	+
E165R	dUTPase	DNA integrity	AIB51748	+	+	+	+
NP419L	DNA ligase	DNA integrity	AIB52048	+	+	+	+
NP1450L	RNA polymerase subunit 1	Transcription	AIB52040	+	+	+	+
EP1242L	RNA polymerase subunit 2	Transcription	AIB51752	+	+	+	+
H359L	RNA polymerase subunit 3-11	Transcription	AIB52132	+	+	+	+
D205R	RNA polymerase subunit 5	Transcription	AIB52137	+		+	+
C147L	RNA polymerase subunit 6	Transcription	AIB51823	+			
D339L	RNA polymerase subunit 7	Transcription	AIB52143	+	+	+	+
Q706L	VACV D11-like helicase	Transcription	AIB52129	+	+	+	+
B962L	VACV I8-like RNA helicase	Transcription	AIB51810	+	+	+	+
D1133L	VACV D6-like RNA helicase	Transcription	AIB52142	+	+	+	+
G1340L	VACV A7 early transcription factor large subunit-like	Transcription	AIB52005	+	+	+	+
NP868R	mRNA-capping enzyme	Transcription	AIB52055	+	+	+	+
C475L	Poly(A) polymerase	Transcription	AIB51816	+			
EP424R	Putative RNA methyltransferase	Transcription	AIB52114	+	+		+
EP152R	Protein EP152R		AIB51785				+
B169L	Uncharacterized protein		AIB51862	+			
H339R	α-NAC binding protein pH339R		AIB52116	+	+	+	+
M1249L	Uncharacterized protein		AIB51842	+	+	+	+
C129R	Uncharacterized protein		AIB51831	+			+
K421R	Uncharacterized protein		AIB51770	+		+	+
H240R	Uncharacterized protein		AIB52125	+		+	+
QP383R	Uncharacterized protein		AIB52104	+	+	+	+
C122R	Uncharacterized protein		AIB51826	+			
M448R	Uncharacterized protein		AIB51841	+	+	+	+

The transcripts corresponding to the proteins of Faustovirus E12 with a homolog in African swine fever virus (ASFV) are indicated with a "+" when detected in the corresponding dataset.

in ASFV purified particles proteome and their expression in late time points. All proteins forming the core shell in ASFV have their homologs in Faustovirus E12 expressed starting from 6 h p.i.: the 220 kDa polyprotein, the 62 kDa polyprotein and the protease necessary for their cleavage into their corresponding products. Other proteins found in the nucleoid of ASFV with their homolog predicted genes being expressed in Faustovirus E12 include all RNA polymerase subunits and RNA modification enzymes, transcription factors and DNA repair enzymes. Interestingly, using sequence homology, we were unable to identify in Faustovirus E12 genes predicted to encode for proteins detected in the outer and inner envelope of ASFV.

Splicing Events in Faustovirus E12

Using a splice-aware mapper and a double-round alignment strategy, with a manual validation of splice-junctions, we were able to identify 26 potential splice-junctions represented by at least two reads, with insert sizes reaching up to 3,256 bp. **Figure 4** illustrates their distribution across the genome and throughout the replicative cycle of Faustovirus E12 in *Vermamoeba vermiformis*. We observed an uneven distribution of potential introns, with a high rate of splice-junctions grouped together in a single region of the genome and appearing in late times p.i. This region is the one predicted in previous studies to encode for the MCP of the virus (Klose et al., 2016). Overall, the number of junction-reads reached 2.7% (1,386) of the total mapped reads with 95.7% (1,326) of these reads aligning to the MCP encoding region.

Faustovirus E12 Major Capsid Protein Transcription

In order to study the transcription of the MCP, we used both the gene prediction results of GeneWise, and the junction-reads after the first-round alignment. Junction-reads confirming the positions of predicted exons were added to the validated junctions file for the second-round alignment. The complete MCP transcript appears composed of 13 exons delineating 12 introns. Nine of the 13 exon–intron boundaries are supported

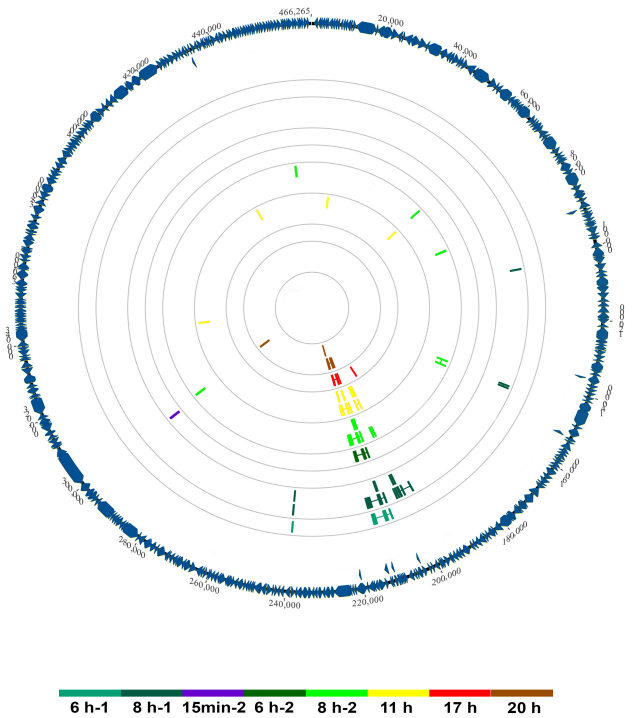


FIGURE 4 | Genome-wide map of splicing events in Faustovirus E12 across its replication cycle. Predicted and curated splice junctions resulting from the second-round mapping of RNA-seq reads against the genome with HISAT2 are reported in *colored boxes* corresponding to the dataset where they were detected. When in close genomic coordinates, junctions appear in two layers for display purpose. The annotated protein coding sequences are represented on the external circle in *blue wedges*.

by detected junction-reads. The mean intron length is 1,273 bp, with minimum and maximum lengths of 396 and 3,256 bp, respectively, and a mean G + C content of 35.2%. Exons forming the MCP coding transcript are significantly shorter (p = 0.0007, unpaired t-test) with length varying from 13 to 527 bp for a mean length of 149 bp and a mean G + C content of 43.9%. The exonic G + C content is significantly higher than that observed in introns ($p < 0.0001$, unpaired t-test).

An A/T substitution at transcript position 1,879 was found to generate a premature stop codon at protein position 631, suggesting the presence of a potential frame shift or post-transcriptional RNA editing mechanism.

Reads mapping to the MCP region represented 23.5% of the total mapped reads. The coverage of the intronic and exonic regions shifts during the replicative cycle. In early time points and until 3 h p.i, Faustovirus E12 appears to express transcripts corresponding to the intronic regions with coverage varying from 1.36 to 3.56, while for the same samples, the exonic regions have a null coverage. Starting from 6 h p.i., the exonic regions are detected and the highest coverage was observed in the sample T11H with 265.12 average coverage versus 16.89 in intronic regions of the same sample.

To get a closer view on the mechanisms involved in the expression of the MCP gene, exon–intron boundaries were examined for conserved splice-sites that would suggest the presence of spliceosome-processed introns. Moreover, the intronic sequences were searched against the Rfam database for conserved motifs. Through this approach, five group I self-splicing introns were identified, and two introns were shown to contain an inserted ORF encoding a GIY-YIG homing endonuclease (**Figure 5A**). All the MCP gene exon-intron boundaries show non-canonical splice-sites with five types of donor-acceptor associations, among which AA/TG was the most represented (**Figure 5B**).

DISCUSSION

An Overview of the Transcriptional Landscape of Faustovirus E12

This study represents the first exploration of the transcriptional landscape of Faustovirus E12. Using total RNA sequencing of infected cells at nine different time points covering the whole replicative cycle of the virus in *Vermamoeba vermiformis*, we were able to follow the temporal regulation of the viral transcription. Faustovirus E12 gene expression seems to follow the classical temporal regulation described in other giant viruses of amoebae and those of the former NCLDV group. Early on, transcripts related to the ubiquitin pathway were detected. This pathway has been described as a viral adaptation mechanism against host defenses. By transcribing its own components of the ubiquitin pathway, the virus can alter the host response to infection by modulating or degrading cell proteins (Iyer et al., 2006). Ankyrin repeats containing proteins are also expressed early and throughout the replicative cycle. In Poxviruses, these motif containing proteins have been described as modulators of host-range and their early expression could play a role in repressing host response by targeting the NF-κB pathway (Herbert et al., 2015). In parallel, to prepare its replication, the virus encodes the ribonucleotide reductase small and large subunits that provide the dNTPs necessary for viral DNA synthesis, therefore allowing virus growth in non-dividing cells (Gammon et al., 2010). Early mRNA transcripts are likely expressed using viral enzymes packaged within the infectious particles. Similarly to what is described in ASFV, the viral RNA polymerase is responsible for the transcription of all the viral genes but is expressed later during the replicative cycle (Rodríguez and Salas, 2013). Indeed, different RNA polymerase subunits, transcription factors and RNA modification enzymes are expressed late during the infection cycle, and likely translated into proteins incorporated to the virions during the assembly step. The comparison of the nucleoid components described by the proteomics analysis of ASFV particles and the late transcribed genes in Faustovirus E12 comforts this hypothesis (Alejo et al., 2018). Among the late transcribed gene products, we identified three enzymes homologous to the components of the base excision repair (BER) pathway described in ASFV. This pathway has been hypothesized to serve as an adaptation mechanism

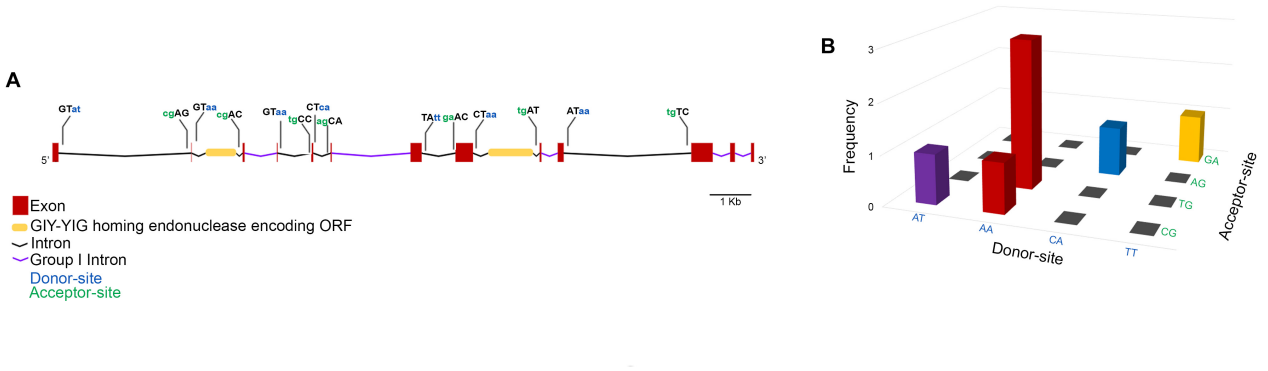


FIGURE 5 | Faustovirus E12 major capsid protein gene structure. The MCP gene contains 13 exons and 12 introns among which five are group I self-spliced introns and two contain a GIY-YIG homing endonuclease inserted ORF. The donor and acceptor splice sites are represented for potentially spliceosome-processed introns **(A)**. The association of donor/acceptor splice site in these introns shows the high abundance of the donor site AA and its frequent association with TG acceptor site **(B)**.

for viral replication in the cytoplasm of macrophages while not expressed in tissue cell cultures (Dixon et al., 2013). Formed by a DNA polymerase type X, a class II Apurinic/apyrimidinic (AP) endonuclease and a DNA ligase, all three detected in the transcriptome of Faustovirus E12 infected *Vermamoeba vermiformis*, this pathway could confirm the potential role of amoebae as training field for microorganisms' resistance to macrophages (Greub and Raoult, 2004). The comparative study of Faustovirus E12 transcription on different host cells should be further investigated.

Faustovirus E12 DNA primase responsible for the initiation of DNA replication (AIB51821) in ASFV and the proliferating cell nuclear antigen-like protein that clamps the DNA polymerase to the DNA (AB52098) are expressed starting from 6 h p.i. with the onset of DNA replication (Dixon et al., 2013; Reteno et al., 2015). The most abundant transcripts detected in our study appear late during the infection cycle after t = 6h, at the viral factory step, and correspond to structural proteins responsible for the particles' morphogenesis and packaging: the MCP, forming the external protein shell, is the most abundant transcript in late times. It is followed by the 220 kDa polyprotein and the 62 kDa polyprotein, both described as essential for the assembly of the core shell and the incorporation of the genomic DNA and nucleoid components in the mature virions (Andrés et al., 2002; Suárez et al., 2010). At t = 20 h p.i. as described in the developmental cycle of Faustovirus E12, most amoebae are lysed (Reteno et al., 2015) or appear at different stages of the replicative cycle of Faustovirus E12. This shows in our data with a mix of early and late transcribed genes in this dataset.

Although our data confirm the expression of most of Faustovirus E12 predicted protein-encoding genes, the low abundance of viral reads doesn't allow further interpretation. With the high abundance of amoebal rRNA and mRNA in the total RNA extract, and in the absence of the *V. vermiformis* complete genome sequence from international sequence databases, the reads that could not be aligned against the Faustovirus E12 genome could not be unequivocally attributed to this amoeba. The possibility of using a ribodepletion strategy should be explored for future transcriptomic studies targeting giant viruses of amoebae.

Corrected MCP Transcript

This study represents a first step forward in the understanding of the non-canonical splicing in Faustovirus E12 MCP expression. The use of paired-end 250 bp-long read sequencing on the MiSeq instrument allowed us to unambiguously identify splice junctions using a splice-aware mapper. Although HISAT2 is adapted to eukaryotic model organisms, the use of both prediction data and manual curation of the junction reads allowed us to describe a 1,939 bp-long transcript generated from a 17 kbp long gene and corresponding to the 645 amino acid-long sequence of the MCP forming the external protein shell of the mature Faustovirus E12 virions.

In early times of the replicative cycle, we observed transcription of regions corresponding to the introns of the MCP gene. Moreover, in the absence of RNA enrichment or selection step in our protocol, the observed transcribed introns in later times could be partially due to the presence of immature pre-mRNA particles in the total RNA extract.

Gene splicing was first described by two teams in Adenovirus 2 in 1977 (Berget et al., 1977; Chow et al., 1977). Subsequently, it has proved extensive in eukaryotes and as a central mechanism in gene regulation and protein diversity generation (Kelemen et al., 2013). The presence of introns has been suggested to increase gene expression by controlling the DNA accessibility or through the regulatory effect of some introns on the RNA polymerase (Hir et al., 2003). In Faustovirus E12, splicing was detected in the MCP gene, a high abundance protein encoding gene, with most spliced reads corresponding to this. This could reinforce the hypothesis that splicing plays a role in increasing gene expression (Alejo et al., 2018). However, the low abundance of viral reads in our datasets was a limit to the high confidence identification of other spliced genes.

Among giant viruses, introns were first described in the MCP encoding gene of *Acanthamoeba polyphaga mimivirus*, the firstly discovered giant virus of amoebae (Azza et al., 2009; Legendre et al., 2010). This gene was depicted as composed of three exons separated by two introns. A recent study compared MCP gene splicing profiles in *Mimiviridae* members from lineages A, B, and C and showed a lineage-independent variation in the structure and synteny of exons and intronic regions of this gene (Boratto et al., 2018). Introns were also detected in other conserved gene from giant viruses of amoebae, including those encoding DNA-dependent RNA polymerases and DNA polymerases (Yoosuf et al., 2012; Philippe et al., 2013; Deeg et al., 2018). Other NCLDV spliced genes include different genes of Paramecium bursaria chlorella virus 1 (PBCV-1) with 2 to 3 different types of introns described: spliceosome processed-like introns are present in the DNA polymerase and the pyrimidine dimer-specific glycosylase (PDG) genes and conserved in different chlorella viruses (Sun et al., 2000; Zhang et al., 2001). Group IB self-splicing introns are reported in a putative transcription factor TFII-like gene (ORF A125L) and in other regions of the viral genome where this intron propagated (Blanc et al., 2014).

In Faustovirus E12, a mixed mechanism may interfere with the expression of the MCP gene: the five group I introns could self-splice while the other exons use non-canonical splice-sites for their excision. The splice-sites, defined by the exon–intron boundaries in this virus are different from the usual canonical splice-sites observed in amoebae and eukaryotic cells, making it difficult to accurately identify them by using existing mapping programs alone. The use of known protein sequence to validate splice-junctions and a two-round alignment approach were beneficial for the definition of the MCP gene structure. The 13 exons forming this gene exhibit higher G + C content than their long flanking introns. This difference in G + C content could therefore play a role in the recognition of the exons by the splicing machinery, lowering the constraint on the intron-defined splice-sites, as hypothesized in higher eukaryotes (Amit et al., 2012).

Moreover, this Faustovirus E12 MCP gene exhibits inserted GIY-YIG homing endonuclease encoding ORFs in two different introns. The presence of this enzyme has been thought to play a role in host competition among related viruses, impeding virus replication by cleaving genes essential to virus replication and contributing to the creation of chimeric genomic regions containing parasitic genetic elements in these genomes (Deeg et al., 2018).

As a summary, Faustovirus E12 MCP splicing presents three main features that make it unusual: (i) The number of introns: although splicing has been described in other viruses, the number of introns is generally limited to 1–3 introns. It is, to our knowledge, the first description of a spliced gene composed of 12 introns in a virus. (ii) The size of introns: with a mean length of 1,273 bp, the introns forming the MCP gene of Faustovirus E12 are larger than previously described introns in viruses. The gene structure with multiple large introns is otherwise common in cellular organisms. (iii) The mixed mechanisms that could be in play in the splicing of the MCP gene: Faustovirus E12 MCP gene is formed both of group I introns, and potential spliceosomal introns. Moreover, the potentially spliceosomal introns use non-canonical splice-sites in their excision. Overall, the complexity and unusual splicing observed in Faustovirus E12 contribute to blurring the border between giant viruses of amoebae and cellular organisms, and thus strengthen the delineation of these viruses as different complex entities compared to classical viruses.

DATA AVAILABILITY

The datasets generated for this study were submitted to the European Nucleotide Archive database and are available under the accession numbers ERR2724024 to ERR2724038.

AUTHOR CONTRIBUTIONS

ACL and BLS conceived and designed the experiments. AL, PC, and EB contributed to materials and analysis tools. ACL, AL, PC, and EB analyzed the data. ACL, AL, PC, and BLS wrote the paper.

FUNDING

This work was supported by a grant from the French State managed by the National Research Agency under the "Investissements d'avenir" (Investments for the Future) program with the reference ANR-10-IAHU-03 (Méditerranée Infection) and Région Provence-Alpes-Côte d'Azur and European funding FEDER PRIMI.

ACKNOWLEDGMENTS

We are thankful to Prof. Christophe Beroud for fruitful discussions about splicing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02534/full#supplementary-material

REFERENCES

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 44, W3–W10. doi: 10.1093/nar/gkw343

Alejo, A., Matamoros, T., Guerra, M., and Andrés, G. (2018). A proteomic atlas of the African swine fever virus particle. *J. Virol.* (in press). doi: 10.1128/JVI.01293-18

Alonso, C., Borca, M., Dixon, L., Revilla, Y., Rodriguez, F., Escribano, J. M., et al. (2018). ICTV virus taxonomy profile: *Asfarviridae*. *J. Gen. Virol.* 99, 10–12. doi: 10.1099/jgv.0.000985

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Amit, M., Donyo, M., Hollander, D., Goren, A., Kim, E., Gelfman, S., et al. (2012). Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep.* 1, 543–556. doi: 10.1016/j.celrep.2012.03.013

Andreani, J., Khalil, J. Y. B., Sevvana, M., Benamar, S., Di Pinto, F., Bitam, I., et al. (2017). Pacmanvirus, a new giant icosahedral virus at the crossroads between *Asfarviridae* and *Faustovirus*. *J. Virol.* 91, e212–e217. doi: 10.1128/JVI.00212-17

Andrés, G., García-Escudero, R., Salas, M. L., and Rodríguez, J. M. (2002). Repression of African swine fever virus polyprotein pp220-encoding gene leads to the assembly of icosahedral core-less particles. *J. Virol.* 76, 2654–2666. doi: 10.1128/JVI.76.6.2654-2666.2002

Azza, S., Cambillau, C., Raoult, D., and Suzan-monti, M. (2009). Revised *Mimivirus* major capsid protein sequence reveals intron-containing gene structure and extra domain. *BMC Mol. Biol.* 10:39. doi: 10.1186/1471-2199-10-39

Bajrai, L. H., Benamar, S., Azhar, E. I., Robert, C., Levasseur, A., Raoult, D., et al. (2016). Kaumoebavirus, a new virus that clusters with *Faustoviruses* and *Asfarviridae*. *Viruses* 8:278. doi: 10.3390/v8110278

Benamar, S., Reteno, D. G. I., Bandaly, V., Labas, N., Raoult, D., and La Scola, B. (2016). *Faustoviruses*: comparative genomics of new megavirales family members. *Front. Microbiol.* 7:3. doi: 10.3389/fmicb.2016.00003

Berget, S. M., Moore, C., and Sharp, P. A. (1977). Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U.S.A.* 74, 3171–3175. doi: 10.1073/PNAS.74.8.3171

Blanc, G., Mozar, M., Agarkova, I. V., Gurnon, J. R., Yanai-Balser, G., Rowe, J. M., et al. (2014). Deep RNA sequencing reveals hidden features and dynamics of early gene transcription in *Paramecium bursaria Chlorella* virus 1. *PLoS One* 9:e90989. doi: 10.1371/journal.pone.0090989

Boratto, P. V. M., Dornas, F. P., da Silva, L. C. F., Rodrigues, R. A. L., Oliveira, G. P., Cortines, J. R., et al. (2018). Analyses of the kroon virus major capsid gene and its transcript highlight a distinct pattern of gene evolution and splicing among mimiviruses. *J. Virol.* 92, e1782–e1717. doi: 10.1128/JVI.01782-17

Cherif Louazani, A., Andreani, J., Ouarhache, M., Aherfi, S., Baptiste, E., Levasseur, A., et al. (2017). Genome sequences of new *Faustovirus* strains st1 and lc9, isolated from the South of France. *Genome Announc.* 5, e613–e617. doi: 10.1128/genomeA.00613-17

Chow, L. T., Gelinas, R. E., Broker, T. R., and Roberts, R. J. (1977). An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. *Cell* 12, 1–8. doi: 10.1016/0092-8674(77)90180-5

Deeg, C. M., Chow, C.-E. T., and Suttle, C. A. (2018). The kinetoplastid-infecting *Bodo saltans* virus (BsV), a window into the most abundant giant viruses in the sea. *eLife* 7:e33014. doi: 10.7554/eLife.33014

Dixon, L. K., Chapman, D. A. G., Netherton, C. L., and Upton, C. (2013). African swine fever virus replication and genomics. *Virus Res.* 173, 3–14. doi: 10.1016/j.virusres.2012.10.020

Gammon, D. B., Gowrishankar, B., Duraffour, S., Andrei, G., Upton, C., and Evans, D. H. (2010). *Vaccinia virus* - encoded ribonucleotide reductase subunits are differentially required for replication and pathogenesis. *PLoS Pathog.* 6:e1000984. doi: 10.1371/journal.ppat.1000984

Greub, G., and Raoult, D. (2004). Microorganisms resistant to free-living Amoebae. *Clin. Microbiol. Rev.* 17, 413–433. doi: 10.1128/CMR.17.2.413-433.2004

Herbert, M. H., Squire, C. J., and Mercer, A. A. (2015). Poxviral ankyrin proteins. *Viruses* 7, 709–738. doi: 10.3390/v7020709

Hir, H. L., Nott, A., and Moore, M. J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* 28, 215–220. doi: 10.1016/S0968-0004(03)00052-5

Iyer, L. M., Aravind, L., and Koonin, E. V. (2001). Common origin of four diverse families of large Eukaryotic DNA viruses. *J. Virol.* 75, 11720–11734. doi: 10.1128/JVI.75.23.11720-11734.2001

Iyer, L. M., Balaji, S., Koonin, E. V., and Aravind, L. (2006). Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* 117, 156–184. doi: 10.1016/j.virusres.2006.01.009

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., et al. (2013). Function of alternative splicing. *Gene* 514, 1–30. doi: 10.1016/j.gene.2012.07.083

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Klasberg, S., Bitard-Feildel, T., and Mallet, L. (2016). Computational identification of novel genes: current and future perspectives. *Bioinforma. Biol. Insights* 10, 121–131. doi: 10.4137/BBI.S39950

Klose, T., Reteno, D. G., Benamar, S., Hollerbach, A., Colson, P., La Scola, B., et al. (2016). Structure of *Faustovirus*, a large dsDNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 113, 6206–6211. doi: 10.1073/pnas.1523999113

Legendre, M., Audic, S., Poirot, O., Hingamp, P., Seltzer, V., Byrne, D., et al. (2010). mRNA deep sequencing reveals 75 new genes and a complex transcriptional landscape in *Mimivirus*. *Genome Res.* 20, 664–674. doi: 10.1101/gr.102582.109

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a *pandoravirus* morphology. *Proc. Natl. Acad. Sci.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (2015). In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting *Acanthamoeba*. *Proc. Natl. Acad. Sci.* 112, E5327–E5335. doi: 10.1073/pnas.1510795112

Legendre, M., Santini, S., Rico, A., Abergel, C., and Claverie, J.-M. (2011). Breaking the 1000-gene barrier for *Mimivirus* using ultra-deep genome and transcriptome sequencing. *Virol. J.* 8:99. doi: 10.1186/1743-422X-8-99

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., et al. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43, W580–W584. doi: 10.1093/nar/gkv279

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

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

Rodríguez, J. M., and Salas, M. L. (2013). African swine fever virus transcription. *Virus Res.* 173, 15–28. doi: 10.1016/j.virusres.2012.09.014

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–864. doi: 10.1093/bioinformatics/btr026

Suárez, C., Salas, M. L., and Rodríguez, J. M. (2010). African swine fever virus polyprotein pp62 is essential for viral core development. *J. Virol.* 84, 176–187. doi: 10.1128/JVI.01858-09

Sun, L., Li, Y., McCullough, A. K., Wood, T. G., Lloyd, R. S., Adams, B., et al. (2000). Intron conservation in a UV-specific DNA repair gene encoded by *Chlorella* viruses. *J. Mol. Evol.* 50, 82–92. doi: 10.1007/s002399910009

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

Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–518. doi: 10.1038/nbt.1621

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484

Yoosuf, N., Yutin, N., Colson, P., Shabalina, S. A., Pagnier, I., Robert, C., et al. (2012). Related giant viruses in distant locations and different habitats: *Acanthamoeba polyphaga* moumouvirus represents a third lineage of the Mimiviridae that is close to the megavirus lineage. *Genome Biol. Evol.* 4, 1324–1330. doi: 10.1093/gbe/evs109

Yutin, N., Wolf, Y. I., and Koonin, E. V. (2014). Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* 46, 38–52. doi: 10.1016/j.virol.2014.06.032

Zhang, Y., Adams, B., Sun, L., Burbank, D. E., and Van Etten, J. L. (2001). Intron conservation in the DNA polymerase gene encoded by *Chlorella* viruses. *Virology* 285, 313–321. doi: 10.1006/viro.2001.0935

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 17 November 2017
doi: 10.3389/fmicb.2017.02196

Suppression of Poxvirus Replication by Resveratrol

*Shuai Cao[1], Susan Realegeno[2], Anil Pant[1], Panayampalli S. Satheshkumar[2] and Zhilong Yang[1]**

[1] Division of Biology, Kansas State University, Manhattan, KS, United States, [2] Poxvirus and Rabies Branch, Division of High-Consequence Pathogens and Pathology, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, United States

Poxviruses continue to cause serious diseases even after eradication of the historically deadly infectious human disease, smallpox. Poxviruses are currently being developed as vaccine vectors and cancer therapeutic agents. Resveratrol is a natural polyphenol stilbenoid found in plants that has been shown to inhibit or enhance replication of a number of viruses, but the effect of resveratrol on poxvirus replication is unknown. In the present study, we found that resveratrol dramatically suppressed the replication of vaccinia virus (VACV), the prototypic member of poxviruses, in various cell types. Resveratrol also significantly reduced the replication of monkeypox virus, a zoonotic virus that is endemic in Western and Central Africa and causes human mortality. The inhibitory effect of resveratrol on poxviruses is independent of VACV N1 protein, a potential resveratrol binding target. Further experiments demonstrated that resveratrol had little effect on VACV early gene expression, while it suppressed VACV DNA synthesis, and subsequently post-replicative gene expression.

Keywords: poxvirus, vaccinia virus, monkeypox, resveratrol, DNA synthesis, gene expression, antiviral

OPEN ACCESS

Edited by:
Jonatas Abrahao, Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Danilo Oliveira, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brazil
Iara Apolinario Borges, Universidade Federal de Minas Gerais, Brazil

***Correspondence:**
Zhilong Yang
zyang@ksu.edu

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

Received: *31 August 2017*
Accepted: *26 October 2017*
Published: *17 November 2017*

Citation:
Cao S, Realegeno S, Pant A, Satheshkumar PS and Yang Z (2017) Suppression of Poxvirus Replication by Resveratrol. Front. Microbiol. 8:2196. doi: 10.3389/fmicb.2017.02196

INTRODUCTION

Smallpox is a deadly disease, responsible for approximately 300 million human deaths in the 20th century alone. Smallpox is caused by the variola virus, the most notorious member of the family *Poxviridae* (Miller et al., 2001). Despite the eradication of smallpox 37 years ago, poxviruses are of renewed interest due to their continuous impact on public health. Specifically, many poxviruses cause other human and animal diseases. For example, monkeypox, a zoonotic disease endemic in Central and Western Africa, caused an outbreak in humans in the United States (US) in 2003 (Reed et al., 2004; Bayer-Garner, 2005). Molluscum contagiosum accounts for 1 in 500 outpatient visits per year in the United States (Reynolds et al., 2009). Additionally, there is a concern that variola virus, the causative agent of smallpox, can potentially be used as a biological weapon from unsecured stocks or genetic engineering. Humans are particularly vulnerable to smallpox in the post-smallpox immunization era due to the absence of routine vaccination, waning immunity, and lower proportion of vaccinated individuals in the current population. In fact, between 1980 and 2010, the monkeypox incidence in Central Africa has increased 20 times after the discontinuation of smallpox immunization (Rimoin et al., 2010). In addition, poxviruses are developed as vectors for vaccine development against infectious diseases and as anti-cancer agents (Rerks-Ngarm et al., 2009; Draper and Heeney, 2010; Breitbach et al., 2011; Altenburg et al., 2014; Izzi et al., 2014). There are no FDA-approved drugs for poxvirus-infection treatment. Cidofovir, a drug for human cytomegalovirus infection, is an off-label drug to treat poxvirus infection (Robbins et al., 2005;

Lu et al., 2011; Dower et al., 2012). There were also a number of small-molecule inhibitors of poxviruses identified in the past years, for example, CMX001, Tecovirimat (ST-246), and CMLDBU6128 (Quenelle et al., 2007; Huggins et al., 2009; Jordan et al., 2009). However, resistant viruses to the compounds were isolated in cell culture, including CMX001 and ST-246 (Yang et al., 2005; Andrei et al., 2006; Farlow et al., 2010). A combination therapy may be required to treat infected individuals, which demands the identification and characterization of additional poxvirus inhibitors.

Resveratrol is a natural polyphenol stilbenoid found in grapes, berries, and a number of other plants. Extensive studies have been carried out to investigate its functions in modulating lifespan, metabolism, cancer, and other diseases (Fremont, 2000). Resveratrol inhibits replication of a number of viruses, such as influenza virus, herpes simplex virus, enterovirus, hepatitis C virus, respiratory syncytial virus, human immunodeficiency virus, varicella zoster virus, Epstein-Barr virus, African swine fever virus, and duck enteritis virus (Docherty et al., 1999, 2006; Palamara et al., 2005; Nakamura et al., 2010; Galindo et al., 2011; Espinoza et al., 2012; Xie et al., 2012; Xu et al., 2013; Abba et al., 2015; Zhang et al., 2015). The antiviral mechanisms of resveratrol against these viral infections are diverse and include inhibition of viral protein synthesis, DNA synthesis, and modulation of host functions important for viral infection (Abba et al., 2015). In contrast to the above-mentioned viruses, resveratrol facilitates Kaposi's-sarcoma associated herpesvirus (KSHV) reactivation from latency in several cell lines through enhancing mitochondrial function of infected cells (Yogev et al., 2014). Nevertheless, the effect of resveratrol on poxvirus replication has not been examined. A previous study showed that several polyphenols, including resveratrol, directly bind to and may inhibit vaccinia virus (VACV, the prototypic member of poxviruses)-encoded N1 protein, a cellular apoptotic regulator (Cheltsov et al., 2010). However, N1L is a non-essential gene and deletion of N1L from VACV genome does not affect VACV infection in cultured cells (Bartlett et al., 2002). Therefore, it is unlikely that resveratrol can prevent VACV infection through N1 protein in cell culture.

Here, we demonstrated that resveratrol could strongly suppress VACV replication in multiple cell types. We also showed that resveratrol directly targeted VACV DNA synthesis step and the suppression was independent of the viral N1 protein. Resveratrol also suppressed monkeypox virus (MPXV) replication.

MATERIALS AND METHODS

Cell Culture

BS-C-1 cells (ATCC-CCL26) were cultured in Eagle's Minimum Essential Medium (EMEM). HeLa cells (ATCC-CCL2) were cultured in Dulbecco's Modified Eagle Medium (DMEM). Normal human dermal fibroblasts (NHDFs, ATCC PCS-201-010) and human foreskin fibroblasts (HFFs, kindly provided by Dr. Bernard Moss) were also cultured in DMEM. The EMEM and DMEM were supplemented with 10% fetal bovine serum (FBS), L-glutamine (2 mM), streptomycin (100 μg/mL), and penicillin (100 units/mL). Cells were cultured in an incubator with 5% CO_2 at 37°C.

Cell Viability Assay and Calculation of 50% Cytotoxicity Concentration (CC_{50})

HeLa cells and HFFs were cultured in 12-well plates. The cells were treated with DMSO or resveratrol at a series of concentrations. Cell viability was measured using trypan-blue exclusion test (Strober, 2015). After 24 h of treatment, cells in each well were treated with 300 μL of trypsin and resuspended with 500 μL of DMEM by pipetting. Twenty microliters of cell suspension was gently mixed with 20 μL of 4% trypan blue. The numbers of cells were measured with a hemocytometer. The CC_{50} was calculated using relative cell viability at different resveratrol concentrations by linear regression analysis.

Viruses, Viral Infection, and Titration

Vaccinia virus Western Reserve (WR, ATCC VR-1354) strain was amplified and purified as described previously (Earl et al., 2001a). Recombinant N1L-deleted VACV was generated by homologous recombination and the N1L gene was replaced with a green fluorescent protein (GFP) gene. Briefly, PCR product of GFP coding sequence under a late P11 promoter flanked by 500-bp homologous sequences upstream and downstream N1L gene was transfected into VACV-infected HeLa cells. The transfected cells were collected at 24 h post-infection (hpi). Recombinant viruses expressing GFP were clonally purified by multiple rounds of plaque isolation (Earl et al., 2001b). Recombinant VACV with the correct insertion or deletion was verified by PCR. The recombinant VACV that expresses GFP under a synthetic early/late VACV promoter (Chakrabarti et al., 1997) and dsRED under P11 VACV promoter was generated using a similar procedure. Recombinant virus vP11-Fluc that expresses firefly luciferase gene under the late VACV P11 promoter was described elsewhere (Bengali et al., 2011). MPXV MPXV-WA 2003-044 and MPXV-ROC 2003-358 clades were utilized in this study. Preparation, infection, and titration of VACV and MPXV were carried out as described previously (Earl et al., 2001a). For infection, cells were incubated with desired amount of viruses in DMEM (containing 2.5% FBS). After 1 h of incubation at 37°C in 5% CO_2, virus-containing DMEM was replaced with fresh DMEM (containing 2.5% FBS) and further incubated for desired amount of time. For titration, BS-C-1 cells cultured in 6- or 12-well plates were infected with serial diluted viral samples and incubated in DMEM (containing 2.5% FBS and 0.5% methyl cellulose) for 48 h. The cells were stained with 0.1% crystal violet for 5 min and washed with water before counting the number of plaques.

Measurement and Calculation of 50% Inhibiting Concentration (IC_{50})

HeLa cells or HFFs were cultured in 12-well plates. The cells were infected with VACV at a multiplicity of infection (MOI) of 1 in the presence of DMSO or resveratrol at a series of

concentrations. After 24 hpi, virus titers were measured by a plaque assay. The IC_{50} was calculated using virus inhibitory efficiency at different resveratrol concentrations by linear regression analysis.

Antibodies and Chemical Inhibitors

Antibodies against VACV L2 protein, P4a (A10) protein, and whole VACV viral particle were kindly provided by Dr. Bernard Moss. Antibody against human GAPDH was purchased from Abcam (Cambridge, MA, United States). Chemicals cytosine-1-β-D-arabinofuranoside (AraC), resveratrol, and hydroxyurea were purchased from Sigma (St. Louis, MO, United States).

Western Blotting Analysis

Cells were collected and lysed in NP-40 cell lysis buffer (150 mM NaCl, 1% NP-40, 50 mM Tris–Cl, pH 8.0). Cell lysates were reduced by 100 mM DTT and denatured by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) loading buffer and boiling for 3 min before SDS–PAGE, followed by transferring to a polyvinylidene difluoride membrane. The membrane was then blocked in TBS-Tween (TBST) [50 mM Tris–HCl (pH 7.5), 200 mM NaCl, 0.05% Tween 20] containing 5% skim milk and 1% bovine serum albumin for 1 h, incubated with primary antibody in the same TBST-milk buffer for 1 h, washed with TBST three times for 10 min each time, incubated with horseradish peroxidase-conjugated secondary antibody for 1 h, washed three times with TBST, and developed with chemiluminescent substrate (National Diagnostics, Atlanta, GA, United States). The whole procedure was carried out at room temperature. Antibodies were stripped from the membrane by Restore (Thermo Fisher Scientific, Waltham, MA, United States) for western blot analysis using another antibody.

Luciferase Assay

Firefly luciferase activities were measured by an ENSPIRE plate reader (PerkinElmer, Waltham, MA, United States) using the Luciferase Assay System (Promega, Madison, WI, United States) according to manufacturer's instructions.

Plasmid Replication in VACV-Infected Cells

Total DNA was isolated using E.Z.N.A.® Blood DNA Kit (Omega Bio-Tek, Inc., Norcross, GA, United States). One microgram of DNA was treated with a DpnI enzyme to digest originally transfected input plasmid DNA (amplified from *Escherichia coli*, with methylation on DpnI recognition site) but not the plasmid DNA amplified in mammalian cells (no methylation in DpnI site). The plasmid DNA amounts were then measured using qPCR using a pair of primers that amplify a fragment containing a DpnI site.

Quantitative Real-Time PCR

Total DNA was extracted from mock- or VACV-infected cells at indicated time points using E.Z.N.A.® Blood DNA Kit. Relative viral DNA levels were quantified by CFX96 real-time PCR instrument (Bio-Rad, Hercules, CA, United States) with All-in-one™ 2× qPCR mix (GeneCopoeia) and primers specific for VACV and human genomes, respectively. The qPCR program was started with initial denaturation step at 95°C for 3 min, followed by 40 cycles of denaturation at 95°C for 10 s, annealing and reading fluorescence at 52°C for 30 s, and extension at 72°C for 30 s. The primers used in this study are:

C11pF: AAACACACACTGAGAAACAGCATAAA;
C11pR: ACTATCGGCGAATGATCTGATTATC;
GAPDH-F: ACATCAAGAAGGTGGTGAAGCA;
GAPDH-R: CTTGACAAAGTGGTCGTTGAGG.
The primers used for recombinant N1L-deletion VACV characterization are:
N1-F: TTATTTTTCACCATATAGATCAATCATTAGATCAT.
N1-R: ATGAGGACTCTACTTATTAGATATATTCTTTGGAG.
Puc19-F: TGCGCGTAATCTGCTGCTTG.
Puc19-R: CGAGGTATGTAGGCGGTGCT.

Statistical Analysis

All titration data were represented as the means of at least three independent experiments. One-tailed paired *T*-test was used to access for significant difference between two means with $P < 0.05$.

RESULTS

Resveratrol Suppresses VACV Replication in Immortal and Primary Human Cells

To test the effect of resveratrol on the viability, HeLa cells, an immortal cervical cancer cell line (Scherer et al., 1953), were treated at a series of concentrations. Cell viability assay showed that resveratrol caused 50% HeLa cell death (CC_{50}) at the concentrations of 157.75 μM in 24 h (**Figure 1A** and **Table 1**). Consistent with the result, no significant morphological change was observed for HeLa cells at the concentration of 50 μM (**Figure 1B**). We then examined the effect of resveratrol on VACV replication. HeLa cells were infected with VACV at an MOI of 1 in the presence of a series of concentrations of resveratrol and the viral titers were measured 24 hpi. The concentration of resveratrol that resulted in 50% inhibition (IC_{50}) of VACV replication was 4.72 μM (**Figure 1C** and **Table 1**). Resveratrol reduced virus yield by more than 120-fold at the concentration of 50 μM (**Figure 1C**). The inhibitory effect of VACV replication by resveratrol is comparable to a well-characterized VACV inhibitor, hydroxyurea, which is known to prevent VACV DNA synthesis and decreased virus yield by approximately 200-fold at the concentration of 10 mM under the same infection conditions (**Figure 1D**). We examined the effect of resveratrol on multiple rounds of VACV replication by infecting HeLa cells at a low MOI of 0.01 and measuring the viral yield at different times post VACV infection. We observed significant reduction of viral titers in resveratrol-treated cells started from 8 hpi (**Figure 1E**), again demonstrating that resveratrol severely

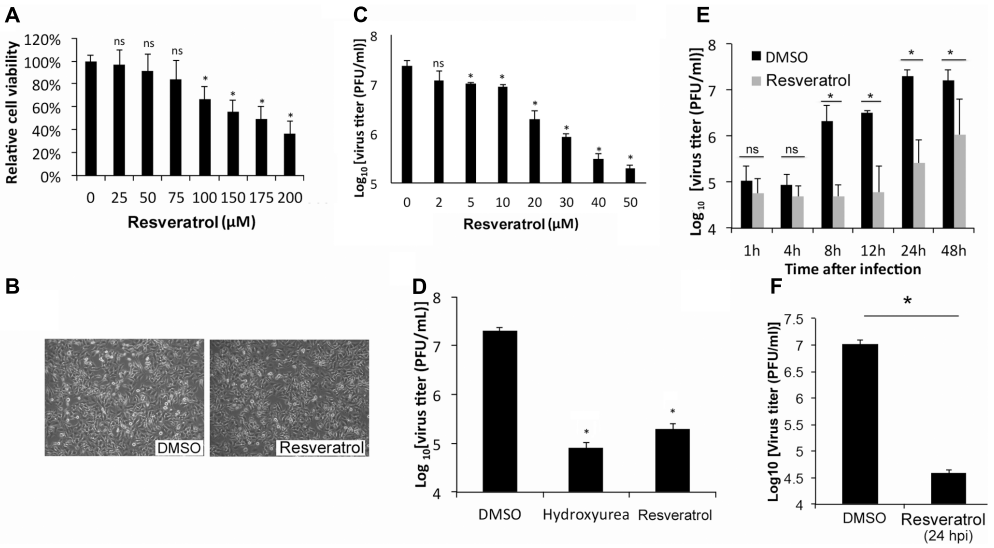


FIGURE 1 | Resveratrol suppresses VACV replication in HeLa cells. **(A)** HeLa cells were treated with DMSO or resveratrol at the indicated concentrations for 24 h. Cell viability was measured using trypan blue exclusion test. **(B)** HeLa cells were imaged under the bright field of a microscope after DMSO or resveratrol (50 μM) treatment for 24 h. **(C)** HeLa cells were infected with VACV at an MOI of 1 in the presence of resveratrol at the indicated concentrations. Virus yield at 24 hpi were determined by a plaque assay. **(D)** HeLa cells were infected with VACV at an MOI of 1 in the presence of DMSO, hydroxyurea (10 mM), or resveratrol (50 μM). Virus yield at 24 hpi were determined by a plaque assay. **(E)** HeLa cells were infected with VACV at an MOI of 0.01 in the presence of DMSO or 50 μM resveratrol. Virus titers were determined by a plaque assay at the indicated time points. **(F)** HeLa cells were infected with VACV at an MOI of 0.001 treated with DMSO or 50 μM resveratrol at 24 hpi. Virus titers were determined by a plaque assay at 72 hpi. The asterisk indicates significant difference ($P < 0.05$) and the ns indicates no significant difference between DMSO-treated cells and resveratrol- or hydroxyurea-treated cells. The error bar indicates standard deviation.

impaired the replication of VACV in HeLa cells. Moreover, the addition of resveratrol at 24 hpi still reduced VACV replication by 250-fold when the initial MOI is low (0.001) (**Figure 1F**), suggesting a possible use of resveratrol to prevent viral spreading post infection.

The effect of resveratrol on VACV replication in primary human cells such as HFFs was also tested. The CC_{50} concentration of resveratrol on HFF was 176.88 μM (**Figure 2A** and **Table 1**). In fact, at the concentration of up to 100 μM, resveratrol did not affect the morphology of HFFs (**Figure 2B**). The IC_{50} concentration of resveratrol in HFFs was 3.51 μM and the virus yield of VACV from 50 μM resveratrol-treated HFFs was reduced by approximately 200-fold at an MOI of 1 (**Figure 2C** and **Table 1**). Moreover, treatment of HFFs with 100 μM of resveratrol protected the HFFs from VACV infection-induced cytopathic effects of the cells (**Figure 2D**). In addition, resveratrol also reduced the replication of VACV in another primary human cell type, NHDF (not shown). Taken together, our results demonstrate that resveratrol dramatically reduces VACV replication in different human cell types.

Resveratrol Suppresses MPXV Replication

We examined the effect of resveratrol on MPXV replication. HeLa cells were infected with MPXV-WA and MPXV-ROC, respectively, at an MOI of 1 in the presence of a series of concentrations of resveratrol and the viral titers were measured 24 hpi. As shown in **Figures 3A,B**, 50 μM resveratrol reduced the virus yield of MPXV-WA and MPXV-ROC clades by 195- and 38-fold, respectively. The IC_{50} was 12.41 μM for WA strain and 15.23 μM for ROC strain (**Table 2**). The inhibitory effect of MPXV replication by resveratrol was comparable to the well-characterized orthopoxvirus (OPXV) inhibitor, AraC, in the corresponding parallel experiments (**Figure 3**).

Resveratrol Suppresses N1L-Deleted VACV Replication

N1L encodes a viral virulence factor that is expressed at early stage of VACV gene expression and regulates host cell apoptosis (Bartlett et al., 2002; Yang et al., 2010). It has been reported that some polyphenols, including resveratrol, could directly bind to and may inhibit the function of N1 protein (Cheltsov et al., 2010). The authors further speculated that resveratrol might inhibit VACV replication by targeting the N1 protein. However, the effect of resveratrol on VACV replication was not tested in the aforementioned study. Moreover, the N1L is not an essential VACV gene and the deletion of N1L from VACV genome was

TABLE 1 | Inhibitory effect of resveratrol on VACV replication and cytotoxicity.

Cells	IC_{50} (μM)[a]	CC_{50} (μM)[b]
HFF	3.51 ± 1.22	176.88 ± 17.44
HeLa	4.72 ± 2.34	157.75 ± 23.66

[a]*The concentration of resveratrol that reduces the yield of VACV by 50%.* [b]*The concentration of resveratrol that causes 50% cell death.*

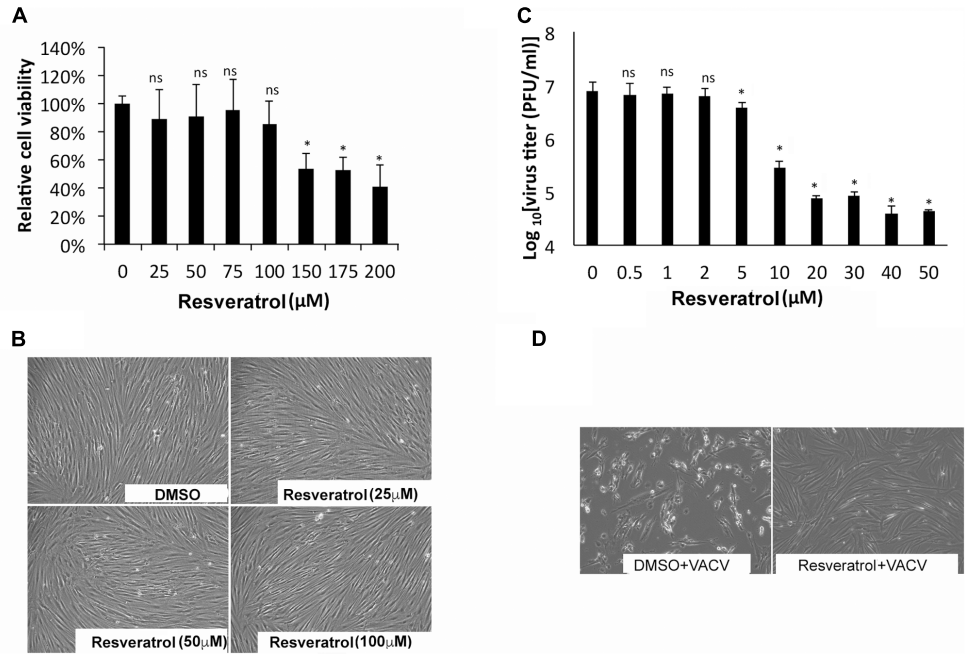


FIGURE 2 | Resveratrol suppresses VACV replication in HFFs. **(A)** HFFs were treated with DMSO or resveratrol at the indicated concentrations for 24 h. Cell viability was measured using trypan blue exclusion test. **(B)** DMSO- or resveratrol-treated cells were imaged under the bright field of a microscope. **(C)** HFFs were infected with VACV at an MOI of 1 in the presence of resveratrol at the indicated concentrations. Virus yield at 24 hpi was determined by a plaque assay. **(D)** HFFs were infected with VACV at an MOI of 0.01 in the presence of resveratrol at the indicated concentrations. At 72 hpi, cells were imaged under the bright field of a microscope. The asterisk indicates significant difference ($P < 0.05$) and the ns indicates no significant difference between DMSO-treated cells and resveratrol-treated cells. The error bar indicates standard deviation.

not shown to affect VACV replication in cultured cells (Bartlett et al., 2002). Based on these facts, we reasoned that prevention of VACV replication by resveratrol is not through the N1 protein. To test it, we replaced the N1L gene with a GFP gene in the VACV genome through homologous recombination (**Figure 4A**). Consistent with a previous observation (Bartlett et al., 2002), the deletion of N1L did not affect VACV replication and viral yields (**Figure 4B**). As expected, resveratrol similarly suppressed VACV-Del-N1L virus (**Figure 4C**), indicating that inhibitory effect is not mediated through the N1 protein.

Resveratrol Suppresses VACV Late, But Not Early Gene Expression

To investigate the stage of viral life cycle targeted by resveratrol, we examined the effect of resveratrol treatment on VACV protein expression by Western blot analysis (**Figure 5A**). Anti-VACV serum was derived from rabbits immunized with purified VACV particles that comprise mostly viral structural proteins expressed at the late stage of VACV gene expression. P4a is a major viral core protein encoded by the VACV late gene A10L (Yang et al., 2011). L2 protein is involved in VACV morphogenesis and is expressed at the early stage of VACV gene expression (Yang et al., 2010; Maruri-Avidal et al., 2011a,b). DNA synthesis inhibitor AraC was used as a positive control. Western blots with anti-VACV serum and P4a antibodies demonstrated dramatic reduction in protein levels in the presence of resveratrol and at levels comparable to the AraC treatment. In contrast, both resveratrol and AraC treatments did not affect the expression level of the viral early protein L2 (**Figure 5A**). We also used a recombinant VACV that expressed GFP under an early/late VACV promoter and dsRED under a late VACV promoter to confirm suppression of late protein synthesis by resveratrol. HeLa cells infected with recombinant VACV expressing fluorescent proteins at an MOI of 1 in the presence of resveratrol, AraC, or vehicle control DMSO were observed under a fluorescent microscope. The results clearly showed that both resveratrol and AraC completely blocked dsRED expression that was expressed at the late stage of gene expression, while they only partially suppressed GFP expression at similar levels that could also be expressed at the early stage of VACV replication (**Figure 5B**). These results indicate that resveratrol has little or only moderate effect on VACV replication prior to viral early gene expression but affects a replication step between the early and late stages of gene expression.

The third approach we employed to examine the effect of resveratrol on VACV late gene expression was using a combination of hydroxyurea and resveratrol. Hydroxyurea blocks VACV DNA synthesis but not early gene expression (Katz et al., 1974). In the control experiment, hydroxyurea and resveratrol were confirmed for their inhibitory effects on expression of VACV late promoter-controlled firefly luciferase gene of vP11-Fluc in HeLa cells (**Figure 5C**). In the parallel experiment, HeLa cells were infected with vP11-Fluc for 3 h in the

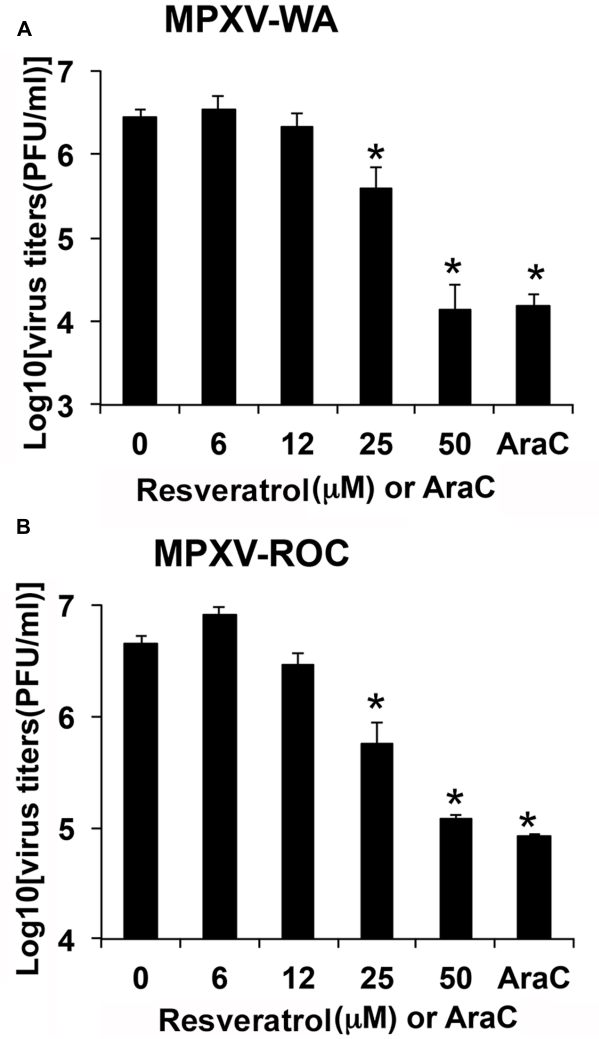


FIGURE 3 | Resveratrol suppresses monkeypox infection. HeLa cells were infected with MPXV-WA **(A)** or MPXV-ROC **(B)** at an MOI of 1 in the presence of resveratrol at the indicated concentrations or AraC (40 μg/mL). Virus yield at 24 hpi was determined by a plaque assay. The asterisk indicates significant difference ($P < 0.05$) between DMSO-treated cells and resveratrol- or AraC-treated cells. The error bar indicates standard deviation.

presence of hydroxyurea, which allowed early gene expression. The hydroxyurea was washed away and DMSO or resveratrol was added and incubated for an additional 3 h. As can be seen, resveratrol still reduced luciferase activity while DMSO could not (**Figure 5C**).

Together, these results indicated that resveratrol affected a post-replication step after VACV early gene expression.

Resveratrol Interferes VACV DNA Synthesis

The effect of resveratrol on VACV DNA synthesis was investigated since it is essential for post-replicative gene expression (intermediate and late protein synthesis). VACV DNA synthesis starts between 2 and 4 hpi in infected HeLa cells under the conditions used in this study (Yang et al., 2010). We examined VACV DNA amounts in VACV-infected HeLa cells at 1 and 24 hpi using quantitative real-time PCR (AraC was used as positive control). Our results indicated that resveratrol treatment significantly reduced VACV DNA amount at 24 hpi (**Figure 6A**). The VACV DNA was 237-fold higher in DMSO-treated cells, while the viral DNA amounts only increased 35- and 6-fold in resveratrol- and AraC-treated cells, respectively (**Figure 6A**).

We tested the direct inhibitory effect on DNA synthesis by resveratrol through examining plasmid DNA synthesis in VACV-infected cells as circular DNA can be replicated in VACV-infected cells that require all known viral proteins needed for VACV DNA synthesis (DeLange and McFadden, 1986; De Silva and Moss, 2005). We transfected pUC19 plasmid into HeLa cells for 12 h and then infected with VACV or mock-infected in the presence or absence of resveratrol and AraC. Total DNA was isolated from cells at 24 hpi and treated with DpnI that only digests methylated input DNA. The DNA was then measured using specific primers amplifying a pUC19 fragment containing the DpnI digestion site. Dpn-resistant plasmid DNA increased 12-fold in VACV-infected cells compared to DMSO treatment. However, there was only a 2- to 3-fold increase of DpnI-resistant plasmid DNA in VACV-infected cells treated with resveratrol or AraC treatment (**Figure 6B**). This result indicated that resveratrol could interfere viral DNA synthesis directly in VACV-infected cells.

DISCUSSION

Our study, for the first time, demonstrated a strong suppressive effect of resveratrol on poxvirus replication. Similar to other viruses, VACV replication is generally divided into entry, gene expression, genome replication, viral particle assembly, and exit steps. VACV gene expression is programmed as a cascade to express viral genes at early, intermediate, and late stages (Moss, 2013a). The early gene expression starts immediately after VACV enters into the infected cells, as the viral infectious particles package all the factors and enzymes needed for early viral mRNA synthesis. The viral early gene products include those necessary factors for viral DNA synthesis. The VACV DNA synthesis is required for viral intermediate, and subsequently, late gene expression. The intermediate and late gene products comprise most of the structural proteins to build infectious viral particles (Moss, 2013a). Our study indicates that the resveratrol directly targets viral DNA synthesis step to prevent VACV replication. Genome uncoating is a step needed to expose encapsidated viral DNA as a template for DNA synthesis. Because resveratrol does not block synthesis of viral early proteins and the viral genome uncoating factor D5 is an early protein (Kilcher et al., 2014),

TABLE 2 | Inhibitory effect of resveratrol on MPXV replication in HeLa cells.

MPXV strains	IC_{50} (μM)[a]
WA	12.41 ± 3.28
ROC	15.23 ± 2.71

[a] *The concentration of resveratrol that reduces the yield of MPXV by 50%.*

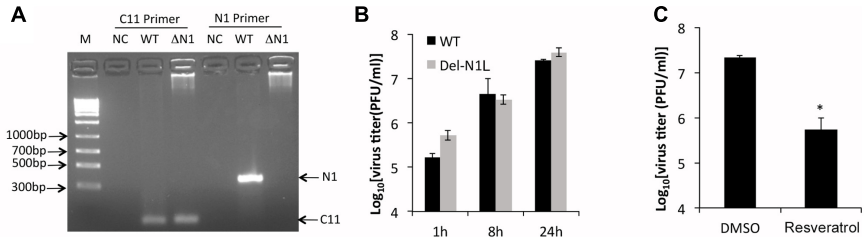


FIGURE 4 | Resveratrol suppresses N1L-deleted VACV replication. **(A)** Deletion of N1L gene from recombinant virus was confirmed by PCR with two pairs of primer specific for C11R gene (positive control) and N1L gene, respectively. **(B)** HeLa cells were infected with wild type and N1L-deleted VACV at an MOI of 1. Virus growth was determined by a plaque assay. **(C)** HeLa cells were infected with N1L-deleted VACV at an MOI of 1 in the presence or absence of resveratrol (50 μM). Virus titers were determined by a plaque assay at 24 hpi. The asterisk indicates significant difference ($P < 0.05$). The error bar indicates standard deviation.

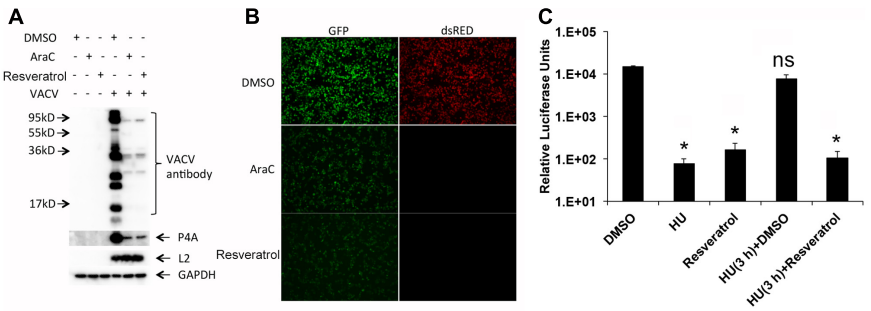


FIGURE 5 | Resveratrol suppresses VACV late, but not early gene expression. **(A)** HeLa cells were infected with VACV at an MOI of 1 in the presence of DMSO, AraC (40 μg/mL), or resveratrol (50 μM). At 24 hpi, the expression of viral proteins in infected cells was detected by western blotting using the indicated antibodies. **(B)** HeLa cells were infected with a recombinant VACV carrying a GFP gene under an early/late VACV promoter and a dsRED gene under a late VACV promoter at an MOI of 1 in the presence of DMSO, AraC (40 μg/mL), or resveratrol (50 μM). The expression of GFP and dsRED was observed and imaged using a fluorescent microscope. **(C)** HeLa cells were infected with VACV-P11-Fluc at an MOI of 1 and treated with hydroxyurea (HU, 10 mM) from 0 to 3 hpi. Then hydroxyurea-containing medium was washed away and replaced with cell culture medium containing DMSO or resveratrol (50 μM), and further incubated for another 3 h until luciferase activity in the infected cell lysates was measured. Luciferase activities from infected cells treated with only DMSO, hydroxyurea, or resveratrol through 0–6 hpi were also measured. The asterisk indicates significant difference ($P < 0.05$) between control and treated cells. The ns indicates no significant difference. The error bar indicates standard deviation.

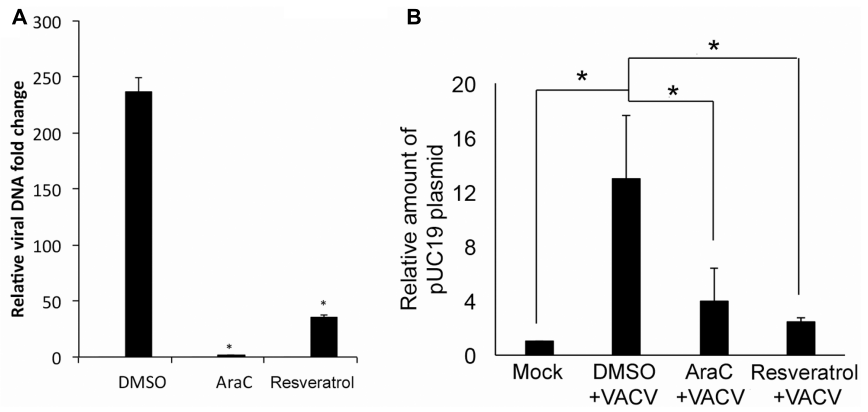


FIGURE 6 | Resveratrol suppresses VACV DNA synthesis. **(A)** HeLa cells were infected with VACV at an MOI of 1 in the presence of DMSO, AraC (40 μg/mL), or resveratrol (50 μM). Relative viral DNA levels in infected cells were determined by real-time PCR at 1 and 24 hpi. The viral DNA level at 24 hpi was determined as the fold to the viral DNA level at 1 hpi. **(B)** HeLa cells were transfected with 200 ng of pUC19 plasmid and incubated overnight. The cells were then infected with VACV at an MOI of 5 or mock-infected in the presence of AraC, resveratrol, or DMSO. Total DNA was extracted from the cells at 24 hpi and 1 μg of total DNA was digested with DpnI at 37°C for 2 h followed by real-time qPCR using primers amplifying pUC19 fragment containing DpnI digestion site. The asterisk indicates significant difference ($P < 0.05$) and the ns indicates no significant difference between 1 and 24 hpi. The error bar indicates standard deviation.

it is unlikely that resveratrol can prevent poxvirus genome uncoating. However, we do not completely rule out the possibility that resveratrol interferes poxvirus genome uncoating to some extent. Interestingly, the effect is independent of the non-essential viral N1L gene, albeit resveratrol has been suggested to be an inhibitor of the VACV N1 protein (Cheltsov et al., 2010).

As the prototypic member of poxvirus family, VACV has a linear, double-stranded DNA genome that replicates entirely in the cytoplasm (Moss, 2013a). The size of the genome is approximately 200 kbp. Although the molecular mechanism involved in VACV DNA synthesis is not fully understood, it is known that the VACV genome encodes most proteins required for replicating its DNA genome (Moss, 2013b). These proteins include a DNA polymerase encoded by E9L gene (Jones and Moss, 1984; Traktman et al., 1984), a helicase–primase encoded by D5R (Roseman and Hruby, 1987), a processivity factor encoded by A20R (McDonald et al., 1997), a uracil DNA glycosylase encoded by D4R (Upton et al., 1993), and a few other proteins bearing different roles in copying the viral DNA (Moss, 2013b). It has been shown that resveratrol inhibits multiple mammalian DNA polymerases including polymerase alpha through its 4-hydroxystyryl moiety, subsequently suppressing active DNA synthesis (Locatelli et al., 2005). As VACV DNA polymerase has considerable similarity to human polymerase alpha (Wang et al., 1989), it is highly possible that resveratrol interferes with VACV DNA polymerase activity directly. Resveratrol also modulates numerous cellular functions (Fremont, 2000); therefore, it is possible that resveratrol affects a cellular function that is important for VACV genome replication. However, the role of cellular functions in VACV DNA synthesis is poorly understood; thus, it is difficult to have an educated prediction of a specific cellular function that may be involved in this process.

All steps of VACV replication, from viral entry and exit, may be targeted for antiviral drug development. For example, mitoxantrone blocks VACV replication by targeting the virion assembly step (Deng et al., 2007). However, the viral DNA synthesis is one of the major targets for anti-poxvirus drug development. Several compounds that are used to treat poxvirus infection target the viral DNA synthesis step. Cidofovir, an acyclic nucleoside that is approved to treat cytomegalovirus infection in AIDS patients also exhibits anti-poxvirus activity by targeting DNA synthesis (Andrei and Snoeck, 2010). The widely used poxvirus inhibitors, AraC, hydroxyurea, and a recently identified inhibitor, CMX001, also target VACV DNA synthesis (Quenelle et al., 2007). The identification of resveratrol as a VACV DNA synthesis inhibitor may allow for developing alternative or compensative strategies to better manage current and re-emergent poxvirus infections and complications caused by poxviruses-based therapeutics.

CONCLUSION

We showed that resveratrol, a member of natural plant polyphenols that is under extensive investigation of its effects on many biological processes, dramatically reduced VACV and MPXV replication. The suppression appears to affect the viral DNA synthesis step. The results will prompt further investigation of its effect on other poxvirus replication steps as well as the mechanism to inhibit VACV replication.

AUTHOR CONTRIBUTIONS

ZY, PS, and SC contributed to the conception of the study. SC, SR, and AP performed the experiments. SC and SR analyzed the data. ZY, SC, and PS wrote the manuscript.

FUNDING

This work was supported, in part by grants from the National Institutes of Health (P20GM113117, project 3) to ZY. AP was also supported by the Johnson Cancer Research Center at Kansas States University.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Bernard Moss at the NIH for providing VACV WR strain, cells, and reagents. The authors also wish to thank other members in the Yang laboratory for helpful discussion. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, Atlanta, GA, United States, and Kansas State University.

REFERENCES

Abba, Y., Hassim, H., Hamzah, H., and Noordin, M. M. (2015). Antiviral activity of resveratrol against human and animal viruses. *Adv. Virol.* 2015:184241. doi: 10.1155/2015/184241

Altenburg, A. F., Kreijtz, J. H., de Vries, R. D., Song, F., Fux, R., Rimmelzwaan, G. F., et al. (2014). Modified *Vaccinia virus* ankara (MVA) as production platform for vaccines against influenza and other viral respiratory diseases. *Viruses* 6, 2735–2761. doi: 10.3390/v6072735

Andrei, G., Gammon, D. B., Fiten, P., De Clercq, E., Opdenakker, G., Snoeck, R., et al. (2006). Cidofovir resistance in *Vaccinia virus* is linked to diminished virulence in mice. *J. Virol.* 80, 9391–9401. doi: 10.1128/JVI.00605-06

Andrei, G., and Snoeck, R. (2010). Cidofovir activity against poxvirus infections. *Viruses* 2, 2803–2830. doi: 10.3390/v2122803

Bartlett, N., Symons, J. A., Tscharke, D. C., and Smith, G. L. (2002). The *Vaccinia virus* N1L protein is an intracellular homodimer that promotes virulence. *J. Gen. Virol.* 83(Pt 8), 1965–1976. doi: 10.1099/0022-1317-83-8-1965

Bayer-Garner, I. B. (2005). Monkeypox virus: histologic, immunohistochemical and electron-microscopic findings. *J. Cutan. Pathol.* 32, 28–34. doi: 10.1111/j.0303-6987.2005.00254.x

Bengali, Z., Satheshkumar, P. S., Yang, Z., Weisberg, A. S., Paran, N., and Moss, B. (2011). Drosophila S2 cells are non-permissive for *Vaccinia virus* DNA replication following entry via low pH-dependent endocytosis and early transcription. *PLOS ONE* 6:e17248. doi: 10.1371/journal.pone.0017248

Breitbach, C. J., Burke, J., Jonker, D., Stephenson, J., Haas, A. R., Chow, L. Q., et al. (2011). Intravenous delivery of a multi-mechanistic cancer-targeted oncolytic poxvirus in humans. *Nature* 477, 99–102. doi: 10.1038/nature10358

Chakrabarti, S., Sisler, J. R., and Moss, B. (1997). Compact, synthetic, *Vaccinia virus* early/late promoter for protein expression. *Biotechniques* 23, 1094–1097.

Cheltsov, A. V., Aoyagi, M., Aleshin, A., Yu, E. C., Gilliland, T., Zhai, D., et al. (2010). *Vaccinia virus* virulence factor N1L is a novel promising target for antiviral therapeutic intervention. *J. Med. Chem.* 53, 3899–3906. doi: 10.1021/jm901446n

De Silva, F. S., and Moss, B. (2005). Origin-independent plasmid replication occurs in *Vaccinia virus* cytoplasmic factories and requires all five known poxvirus replication factors. *Virol. J.* 2:23.

DeLange, A. M., and McFadden, G. (1986). Sequence-nonspecific replication of transfected plasmid DNA in poxvirus-infected cells. *Proc. Natl. Acad. Sci. USA.* 83, 614–618. doi: 10.1073/pnas.83.3.614

Deng, L., Dai, P., Ciro, A., Smee, D. F., Djaballah, H., and Shuman, S. (2007). Identification of novel antipoxviral agents: mitoxantrone inhibits *Vaccinia virus* replication by blocking virion assembly. *J. Virol.* 81, 13392–13402. doi: 10.1128/JVI.00770-07

Docherty, J. J., Fu, M. M., Stiffler, B. S., Limperos, R. J., Pokabla, C. M., and DeLucia, A. L. (1999). Resveratrol inhibition of herpes simplex virus replication. *Antiviral Res.* 43, 145–155. doi: 10.1016/S0166-3542(99)00042-X

Docherty, J. J., Sweet, T. J., Bailey, E., Faith, S. A., and Booth, T. (2006). Resveratrol inhibition of varicella-zoster virus replication in vitro. *Antiviral Res.* 72, 171–177. doi: 10.1016/j.antiviral.2006.07.004

Dower, K., Filone, C. M., Hodges, E. N., Bjornson, Z. B., Rubins, K. H., Brown, L. E., et al. (2012). Identification of a pyridopyrimidinone inhibitor of orthopoxviruses from a diversity-oriented synthesis library. *J. Virol.* 86, 2632–2640. doi: 10.1128/JVI.05416-11

Draper, S. J., and Heeney, J. L. (2010). Viruses as vaccine vectors for infectious diseases and cancer. *Nat. Rev. Microbiol.* 8, 62–73. doi: 10.1038/nrmicro2240

Earl, P. L., Cooper, N., Wyatt, L. S., Moss, B., and Carroll, M. W. (2001a). Preparation of cell cultures and *Vaccinia virus* stocks. *Curr. Protoc. Mol. Biol.* 16, 16.16.1–16.16.18.

Earl, P. L., Moss, B., Wyatt, L. S., and Carroll, M. W. (2001b). Generation of recombinant vaccinia viruses. *Curr. Protoc. Mol. Biol.* 43, 16.17.1–16.17.19.

Espinoza, J. L., Takami, A., Trung, L. Q., Kato, S., and Nakao, S. (2012). Resveratrol prevents EBV transformation and inhibits the outgrowth of EBV-immortalized human B cells. *PLOS ONE* 7:e51306. doi: 10.1371/journal.pone.0051306

Farlow, J., Ichou, M. A., Huggins, J., and Ibrahim, S. (2010). Comparative whole genome sequence analysis of wild-type and cidofovir-resistant monkeypoxvirus. *Virol. J.* 7:110. doi: 10.1186/1743-422X-7-110

Fremont, L. (2000). Biological effects of resveratrol. *Life Sci.* 66, 663–673. doi: 10.1016/S0024-3205(99)00410-5

Galindo, I., Hernaez, B., Berna, J., Fenoll, J., Cenis, J. L., Escribano, J. M., et al. (2011). Comparative inhibitory activity of the stilbenes resveratrol and oxyresveratrol on African swine fever virus replication. *Antiviral Res.* 91, 57–63. doi: 10.1016/j.antiviral.2011.04.013

Huggins, J., Goff, A., Hensley, L., Mucker, E., Shamblin, J., Wlazlowski, C., et al. (2009). Nonhuman primates are protected from smallpox virus or monkeypox virus challenges by the antiviral drug ST-246. *Antimicrob. Agents Chemother.* 53, 2620–2625. doi: 10.1128/AAC.00021-09

Izzi, V., Buler, M., Masuelli, L., Giganti, M. G., Modesti, A., and Bei, R. (2014). Poxvirus-based vaccines for cancer immunotherapy: new insights from combined cytokines/co-stimulatory molecules delivery and "uncommon" strains. *Anticancer Agents Med. Chem.* 14, 183–189. doi: 10.2174/18715206113136660376

Jones, E. V., and Moss, B. (1984). Mapping of the *Vaccinia virus* DNA polymerase gene by marker rescue and cell-free translation of selected RNA. *J. Virol.* 49, 72–77.

Jordan, R., Goff, A., Frimm, A., Corrado, M. L., Hensley, L. E., Byrd, C. M., et al. (2009). ST-246 antiviral efficacy in a nonhuman primate monkeypox model: determination of the minimal effective dose and human dose justification. *Antimicrob. Agents Chemother.* 53, 1817–1822. doi: 10.1128/AAC.01596-08

Katz, E., Margalith, E., and Winer, B. (1974). Synthesis of *Vaccinia virus* polypeptides in the presence of hydroxyurea. *Antimicrob. Agents Chemother.* 6, 647–650. doi: 10.1128/AAC.6.5.647

Kilcher, S., Schmidt, F. I., Schneider, C., Kopf, M., Helenius, A., and Mercer, J. (2014). siRNA screen of early poxvirus genes identifies the AAA+ ATPase D5 as the virus genome-uncoating factor. *Cell Host Microbe* 15, 103–112. doi: 10.1016/j.chom.2013.12.008

Locatelli, G. A., Savio, M., Forti, L., Shevelev, I., Ramadan, K., Stivala, L. A., et al. (2005). Inhibition of mammalian DNA polymerases by resveratrol: mechanism and structural determinants. *Biochem. J.* 389(Pt 2), 259–268. doi: 10.1042/BJ20050094

Lu, Y. C., Fan, H. C., Wang, C. C., and Cheng, S. N. (2011). Concomitant use of acyclovir and intravenous immunoglobulin rescues an immunocompromised child with disseminated varicella caused multiple organ failure. *J. Pediatr. Hematol. Oncol.* 33, e350–e351. doi: 10.1097/MPH.0b013e3181ec0efb

Maruri-Avidal, L., Domi, A., Weisberg, A. S., and Moss, B. (2011a). Participation of *Vaccinia virus* l2 protein in the formation of crescent membranes and immature virions. *J. Virol.* 85, 2504–2511. doi: 10.1128/JVI.02505-10

Maruri-Avidal, L., Weisberg, A. S., and Moss, B. (2011b). *Vaccinia virus* L2 protein associates with the endoplasmic reticulum near the growing edge of crescent precursors of immature virions and stabilizes a subset of viral membrane proteins. *J. Virol.* 85, 12431–12441. doi: 10.1128/JVI.05573-11

McDonald, W. F., Klemperer, N., and Traktman, P. (1997). Characterization of a processive form of the *Vaccinia virus* DNA polymerase. *Virology* 234, 168–175. doi: 10.1006/viro.1997.8639

Miller, J., Engelberg, S., and Broad, W. J. (2001). *Germs: The Ultimate Weapon*. New York, NY: Simon & Schuster Ltd.

Moss, B. (2013a). "Poxviridae: the viruses and their replication," in *Fields Virology*, Vol. 2, eds D. M. Knipe and P. M. Howley (Philadelphia: PA: Wolters Kluwer/Lippincott Williams & Wilkins), 2129–2159.

Moss, B. (2013b). Poxvirus DNA replication. *Cold Spring Harb. Perspect. Biol.* 5, a010199. doi: 10.1101/cshperspect.a010199

Nakamura, M., Saito, H., Ikeda, M., Hokari, R., Kato, N., Hibi, T., et al. (2010). An antioxidant resveratrol significantly enhanced replication of hepatitis C virus. *World J. Gastroenterol.* 16, 184–192. doi: 10.3748/wjg.v16.i2.184

Palamara, A. T., Nencioni, L., Aquilano, K., De Chiara, G., Hernandez, L., Cozzolino, F., et al. (2005). Inhibition of influenza A virus replication by resveratrol. *J. Infect. Dis.* 191, 1719–1729. doi: 10.1086/429694

Quenelle, D. C., Prichard, M. N., Keith, K. A., Hruby, D. E., Jordan, R., Painter, G. R., et al. (2007). Synergistic efficacy of the combination of ST-246 with CMX001 against orthopoxviruses. *Antimicrob. Agents Chemother.* 51, 4118–4124. doi: 10.1128/AAC.00762-07

Reed, K. D., Melski, J. W., Graham, M. B., Regnery, R. L., Sotir, M. J., Wegner, M. V., et al. (2004). The detection of monkeypox in humans in the Western Hemisphere. *N. Engl. J. Med.* 350, 342–350. doi: 10.1056/NEJMoa032299

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., et al. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N. Engl. J. Med.* 361, 2209–2220. doi: 10.1056/NEJMoa0908492

Reynolds, M. G., Holman, R. C., Yorita Christensen, K. L., Cheek, J. E., and Damon, I. K. (2009). The Incidence of Molluscum contagiosum among American Indians and Alaska Natives. *PLOS ONE* 4:e5255. doi: 10.1371/journal.pone.0005255

Rimoin, A. W., Mulembakani, P. M., Johnston, S. C., Lloyd Smith, J. O., Kisalu, N. K., Kinkela, T. L., et al. (2010). Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic Republic of Congo. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16262–16267. doi: 10.1073/pnas.1005769107

Robbins, S. J., Jackson, R. J., Fenner, F., Beaton, S., Medveczky, J., Ramshaw, I. A., et al. (2005). The efficacy of cidofovir treatment of mice infected with ectromelia (mousepox) virus encoding interleukin-4. *Antiviral Res.* 66, 1–7. doi: 10.1016/j.antiviral.2004.12.003

Roseman, N. A., and Hruby, D. E. (1987). Nucleotide sequence and transcript organization of a region of the *Vaccinia virus* genome which encodes a constitutively expressed gene required for DNA replication. *J. Virol.* 61, 1398–1406.

Scherer, W. F., Syverton, J. T., and Gey, G. O. (1953). Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J. Exp. Med.* 97, 695–710. doi: 10.1084/jem.97.5.695

Strober, W. (2015). Trypan blue exclusion test of cell viability. *Curr. Protoc. Immunol.* 111, A.3B.1–A.3B.3. doi: 10.1002/0471142735.ima03bs111

Traktman, P., Sridhar, P., Condit, R. C., and Roberts, B. E. (1984). Transcriptional mapping of the DNA polymerase gene of *Vaccinia virus*. *J. Virol.* 49, 125–131.

Upton, C., Stuart, D. T., and McFadden, G. (1993). Identification of a poxvirus gene encoding a uracil DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A.* 90, 4518–4522. doi: 10.1073/pnas.90.10.4518

Wang, T. S., Wong, S. W., and Korn, D. (1989). Human DNA polymerase alpha: predicted functional domains and relationships with viral DNA polymerases. *FASEB J.* 3, 14–21.

Xie, X. H., Zang, N., Li, S. M., Wang, L. J., Deng, Y., He, Y., et al. (2012). Resveratrol Inhibits respiratory syncytial virus-induced IL-6 production, decreases viral replication, and downregulates TRIF expression in airway epithelial cells. *Inflammation* 35, 1392–1401. doi: 10.1007/s10753-012-9452-7

Xu, J., Yin, Z., Li, L., Cheng, A., Jia, R., Song, X., et al. (2013). Inhibitory effect of resveratrol against duck enteritis virus in vitro. *PLOS ONE* 8:e65213. doi: 10.1371/journal.pone.0065213

Yang, G., Pevear, D. C., Davies, M. H., Collett, M. S., Bailey, T., Rippen, S., et al. (2005). An orally bioavailable antipoxvirus compound (ST-246) inhibits extracellular virus formation and protects mice from lethal orthopoxvirus Challenge. *J. Virol.* 79, 13139–13149. doi: 10.1128/JVI.79.20.13139-13149.2005

Yang, Z., Bruno, D. P., Martens, C. A., Porcella, S. F., and Moss, B. (2010). Simultaneous high-resolution analysis of *Vaccinia virus* and host cell transcriptomes by deep RNA sequencing. *Proc. Natl. Acad. Sci. USA.* 107, 11513–11518. doi: 10.1073/pnas.1006594107

Yang, Z., Reynolds, S. E., Martens, C. A., Bruno, D. P., Porcella, S. F., and Moss, B. (2011). Expression profiling of the intermediate and late stages of poxvirus replication. *J. Virol.* 85, 9899–9908. doi: 10.1128/JVI.05446-11

Yogev, O., Lagos, D., Enver, T., and Boshoff, C. (2014). Kaposi's sarcoma herpesvirus microRNAs induce metabolic transformation of infected cells. *PLOS Pathog.* 10:e1004400. doi: 10.1371/journal.ppat.1004400

Zhang, L., Li, Y., Gu, Z., Wang, Y., Shi, M., Ji, Y., et al. (2015). Resveratrol inhibits enterovirus 71 replication and pro-inflammatory cytokine secretion in rhabdosarcoma cells through blocking IKKs/NF-kappaB signaling pathway. *PLOS ONE* 10:e0116879. doi: 10.1371/journal.pone.0116879

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

The reviewer IAB and handling Editor declared their shared affiliation.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 19 December 2017
doi: 10.3389/fmicb.2017.02539

The *in Vitro* Inhibitory Effect of Ectromelia Virus Infection on Innate and Adaptive Immune Properties of GM-CSF-Derived Bone Marrow Cells Is Mouse Strain-Independent

Lidia Szulc-Dąbrowska[1], Justyna Struzik[1], Joanna Cymerys[1], Anna Winnicka[2], Zuzanna Nowak[3], Felix N. Toka[4] and Małgorzata Gieryńska[1]*

[1] Department of Preclinical Sciences, Faculty of Veterinary Medicine, Warsaw University of Life Sciences, Warsaw, Poland, [2] Department of Pathology and Veterinary Diagnostics, Faculty of Veterinary Medicine, Warsaw University of Life Sciences, Warsaw, Poland, [3] Department of Genetics and Animal Breeding, Faculty of Animal Sciences, Warsaw University of Life Sciences, Warsaw, Poland, [4] Department of Biomedical Sciences, Ross University School of Veterinary Medicine, Basseterre, Saint Kitts and Nevis

OPEN ACCESS

Edited by:
Jonatas Abrahao, Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Jonas Dutra Albarnaz, University of Cambridge, United Kingdom
Rodrigo Araújo Lima Rodrigues, Universidade Federal de Minas Gerais, Brazil

***Correspondence:**
Lidia Szulc-Dąbrowska
lidia_szulc@sggw.pl

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

Received: *29 September 2017*
Accepted: *06 December 2017*
Published: *19 December 2017*

Citation:
Szulc-Dąbrowska L, Struzik J, Cymerys J, Winnicka A, Nowak Z, Toka FN and Gieryńska M (2017) The in Vitro Inhibitory Effect of Ectromelia Virus Infection on Innate and Adaptive Immune Properties of GM-CSF-Derived Bone Marrow Cells Is Mouse Strain-Independent. Front. Microbiol. 8:2539. doi: 10.3389/fmicb.2017.02539

Ectromelia virus (ECTV) belongs to the *Orthopoxvirus* genus of the *Poxviridae* family and is a natural pathogen of mice. Certain strains of mice are highly susceptible to ECTV infection and develop mousepox, a lethal disease similar to smallpox of humans caused by variola virus. Currently, the mousepox model is one of the available small animal models for investigating pathogenesis of generalized viral infections. Resistance and susceptibility to ECTV infection in mice are controlled by many genetic factors and are associated with multiple mechanisms of immune response, including preferential polarization of T helper (Th) immune response toward Th1 (protective) or Th2 (non-protective) profile. We hypothesized that viral-induced inhibitory effects on immune properties of conventional dendritic cells (cDCs) are more pronounced in ECTV-susceptible than in resistant mouse strains. To this extent, we confronted the cDCs from resistant (C57BL/6) and susceptible (BALB/c) mice with ECTV, regarding their reactivity and potential to drive T cell responses following infection. Our results showed that *in vitro* infection of granulocyte-macrophage colony-stimulating factor-derived bone marrow cells (GM-BM—comprised of cDCs and macrophages) from C57BL/6 and BALB/c mice similarly down-regulated multiple genes engaged in DC innate and adaptive immune functions, including antigen uptake, processing and presentation, chemokines and cytokines synthesis, and signal transduction. On the contrary, ECTV infection up-regulated *Il10* in GM-BM derived from both strains of mice. Moreover, ECTV similarly inhibited surface expression of major histocompatibility complex and costimulatory molecules on GM-BM, explaining the inability of the cells to attain full maturation after Toll-like receptor (TLR)4 agonist treatment. Additionally, cells from both strains of mice failed to produce cytokines and chemokines engaged in T cell priming and Th1/Th2 polarization after TLR4 stimulation. These data strongly suggest that *in vitro* modulation of GM-BM innate and adaptive immune functions by ECTV occurs irrespective of

whether the mouse strain is susceptible or resistant to infection. Moreover, ECTV limits the GM-BM (including cDCs) capacity to stimulate protective Th1 immune response. We cannot exclude that this may be an important factor in the generation of non-protective Th2 immune response in susceptible BALB/c mice *in vivo*.

Keywords: ectromelia virus, conventional dendritic cells, Th polarization, immunosuppression, viral evasion strategies

INTRODUCTION

The poxviruses are large DNA viruses that are undoubtedly masters of immune evasion, and have evolved to modulate and inhibit the host immune and inflammatory responses. The poxvirus genome encodes multiple classes of immunomodulatory proteins that act either intracellularly (virostealth and virotransducers) or extracellularly (viromimetics: virokines and viroceptors) and antagonize or compete with molecules critically involved in the host antiviral response. In fact, due to diversity of these genes, a single immunomodulatory protein that is shared by all poxviruses has not been identified yet. Moreover, a unique repertoire of immunomodulatory proteins encoded by each virus species allows it to successfully evade the immune response and survive in its natural host (Stanford et al., 2007). Highly host-specific survival strategy is employed especially by members of the *Orthopoxvirus* genus that exhibit a narrow host range and co-evolved with their natural host, e.g., variola virus (VARV, the causative agent of smallpox) in human and ectromelia virus (ECTV, the causative agent of mousepox) in mice. Meantime, other orthopoxviruses, such as vaccinia (VACV), monkeypox (MPXV), and cowpox (CPXV) viruses, which have a broad host range, are able to infect many different mammalian species and may contribute to the unpredictable outcome of infection in a new host species, e.g., MPXV in humans (McCollum and Damon, 2014). Therefore, a better understanding of the immunomodulatory mechanisms used by orthopoxviruses in their natural hosts is especially important for a full knowledge of their immune evasion strategies employed to control the host immune system.

The mousepox model is an excellent small animal model to study pathogenesis of smallpox, a disease that, despite being eradicated from the globe, now represents one of the most dangerous bioterrorism threats to human society. Smallpox is considered by Centers for Disease Control and Prevention (CDC) in Atlanta as a category A bioterrorism agent due to its easy dissemination, transmission from person to person, and high mortality rates (Riedel, 2005). ECTV shares with VARV several common properties, including: narrow host range and co-evolution with the natural host, high infectivity at low dose, and viral transmission and replication. Moreover, both viruses cause severe diseases with similar pathogenesis, aspects of pathology and immune response, and high mortality rates (Stanford et al., 2007). Therefore, mousepox model is extensively used to study basic questions in immune response regulation during generalized viral infections to eventually develop new prophylactic and therapeutic treatments against orthopoxviruses (Parker et al., 2010).

Within inbred strains of mice there is a genetically determined resistance to severe mousepox. C57BL/6 [H-2^b] mice are resistant to the lethal form of disease, whereas BALB/c [H-2^d] mice are fully susceptible to ECTV infection and usually succumb to disease between 7 and 9 days post-footpad infection. Genetic resistance is controlled by at least four autosomal dominant genes called *rmp* (resistance to mousepox), which are involved in regulation of some aspects of innate immunity (Brownstein and Gras, 1995). Additionally, during the infection resistant and susceptible strains of mice develop different types of the T helper (Th) cytokine immune responses: C57BL/6 mice generate a protective Th1 immune response accompanied by strong cytotoxic T lymphocyte (CTL) activity, whereas BALB/c mice generate a non-protective Th2 immune response, which is associated with a weak/absent CTL activity (Chaudhri et al., 2004).

The central role in driving T cell responses is played by dendritic cells (DCs), the most potent antigen presenting cells (APCs). Depending on lineage, maturation stage, and activation status, DCs release different polarizing signals, the most important of which are cytokines and chemokines that selectively promote the generation of Th1, Th2, Th17, or regulatory T cells (Tregs) (Kaiko et al., 2008). Immature DCs, characterized by low expression of antigen presenting [major histocompatibility complex (MHC) I, MHC II, CD1d] and costimulatory (CD80, CD86, and CD40) molecules induce anergy in antigen-specific naïve T cells or generate $Foxp3^+$ induced Tregs in the presence of transforming growth factor (TGF)-β. Additionally, immature or semi-mature DCs are able to induce regulatory $Foxp3^-$ IL-10^+ Tr1 (type 1 regulatory T) cells. Th2-polarizing DCs have semi-mature state associated with increased expression of antigen presenting and costimulatory molecules, and inability to secrete polarizing cytokines, such as IL-12p70 (Th1) or IL-6 and IL-23 (Th17). Moreover, IL-10 produced by DCs has been associated with propagation of Th2 immunity, however, this cytokine preferentially blocks DC maturation and induces anergy or Tr1 cells. Fully matured DCs, upon lipopolysaccharide (LPS) or CpG oligonucleotide treatment, produce IL-12p70 and possess strong capacity to polarize T cells toward Th1 profile (Lutz, 2016).

As masters of immune evasion, orthopoxviruses are able to control DC functions important for induction of an antiviral immune response. In general, they inhibit numerous functions of these cells, including antigen uptake and presentation, maturation, pro-inflammatory response, and capacity to activate T cells (Engelmayer et al., 1999; Jenne et al., 2000; Hansen et al., 2011; Szulc-Dąbrowska et al., 2017). However, ECTV,

unlike VACV and CPXV, can productively infect conventional DCs (cDCs) *in vitro* (Szulc-Dąbrowska et al., 2017) and *in vivo* (Sei et al., 2015), what indicates strong adaptation capacity of ECTV to the natural host immune cells. It has been shown that cDCs derived from resistant C57BL/6 and susceptible BALB/c mouse strains may differentially react during viral (Pejawar et al., 2005) and bacterial (Jiang et al., 2010) infections *in vitro*. Upon infection, cDCs from C57BL/6 mice underwent higher functional maturation and/or were able to stimulate more potent $CD8^+$ T cell response than cells from BALB/c mice (Pejawar et al., 2005; Jiang et al., 2010).

We therefore confronted the cDCs from resistant (C57BL/6) and susceptible (BALB/c) mice regarding inter-strain differences in reactivity and potential to drive T cell responses following *in vitro* infection with ECTV. Our results showed that ECTV similarly affects innate and adaptive immune functions of granulocyte-macrophage colony-stimulating factor (GM-CSF)-derived bone marrow cells (GM-BM) obtained from both mouse strains. GM-BM infected with ECTV exhibited a profound down-regulation in expression of many genes involved in maturation and activation of DCs, with the exception of *Il10*, which was up-regulated in a strain independent manner. Moreover, ECTV impaired production of chemokines and cytokines engaged in regulation of Th1 and Th2 immune responses, as well as decreased maturation marker expression on the surface of C57BL/6 and BALB/c GM-BM. Collectively, our data suggest that ECTV-infected GM-BM have strong potential to silence the Th immune response, independently of their genetic background.

MATERIALS AND METHODS

Animals

Male C57BL/6 (H-2^b) and BALB/c (H-2^d) mice (8–12 weeks old) were purchased from the animal facility at Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology in Warsaw, Poland. After arrival, animals were kept for 7 days for acclimatization in the animal facility at the Faculty of Veterinary Medicine under controlled temperature and humidity with free access to food and water. Experimental procedures on animals were approved by the 3rd Ethical Committee for Animal Experimentation at Warsaw University of Life Sciences—SGGW (permission no. 34/2012) and were performed in accordance with institutional Guidelines for Care and Use of Laboratory Animals.

Virus

The Moscow strain of ECTV (ATCC, VR-1374) was obtained from American Type Culture Collection (Manassas, VA, United States). The virus was propagated on African green monkey kidney (Vero) cells (ATCC, CCL-81) maintained in DMEM high glucose (HyClone, Logan, UT, United States) supplemented with 5% fetal bovine serum (FBS; HyClone) and 1% antibiotic–antimycotic solution (100 U/ml penicillin, 100 μg/ml streptomycin, and 0.25 μg/ml amphotericin B; Sigma–Aldrich, St. Louis, MO, United States). The virus stock was purified by sucrose cushion centrifugation. Briefly, virus suspension was layered onto 36% sucrose in 1 mM Tris pH 9.0 and centrifuged at 30,000 × *g* for 60 min at 4°C. After purification, virus stock infectivity was determined by plaque formation assay (PFU/ml) on Vero cell monolayer. Inactivation of the virus was performed by a 30 min exposure to UV radiation (120 W, 320 nm) at a working distance of 12 cm from the UV lamp. The absence of plaque formation in the Vero cell monolayer indicated complete inactivation of the virus.

Generation and Enrichment of GM-BM

The protocol used to generate GM-BM was similar to that used by Lutz et al. (1999). Bone marrow was flushed with cold RPMI-1640 medium from femurs and tibias of mice and erythrocytes were removed using ammonium chloride buffer. Cells were then washed and plated at 1×10^6/well in a six-well plate in RPMI-1640 medium containing 10% heat-inactivated FBS, 1% antibiotic solution (100 U/ml penicillin and 100 mg/ml streptomycin; Sigma–Aldrich), 50 μM 2-mercaptoethanol (Sigma–Aldrich), and 20 ng/ml recombinant mouse (rm) GM-CSF (R&D Systems, Minneapolis, MN, United States). Fresh medium with 20 ng/ml rmGM-CSF was added at day 3 of culture and then was partially replaced every 2 days. On day 8 post-culture, cDCs were enriched using MACS $CD11c^+$ labeled magnetic beads (Miltenyi Biotec, Auburn, CA, United States). After MACS separation GM-BM cultures were evaluated for surface expression of cellular markers using the following monoclonal antibodies (mAbs): anti-mouse CD11c-BV421 (N418, Armenian Hamster IgG2; BioLegend, San Diego, CA, United States), anti-CD11b-BV605 (M1/70, rat IgG2b), anti-I-A/I-E-BV711 (M5/114.15.2, rat IgG2b; both from BD Biosciences, San Jose, CA, United States) and anti-CD205-PE-Cy7 (NLDC-145, rat IgG2a; BioLegend) (**Supplementary Figure S1**).

GM-BM Infection and Treatment

MACS-separated $CD11c^+$ cells (purity ≥95%) were infected with live ECTV at multiplicity of infection (MOI) of 1. In some experiments, cells were exposed to UV-inactivated ECTV (uvi-ECTV) at MOI of 1 (before inactivation). As a control, non-infected cells were cultured in parallel in complete RPMI-1640 medium. Additionally, non-, ECTV-, and/or uvi-ECTV-exposed cells were left untreated or were treated for 24 h with 1 μg/ml LPS (*Escherichia coli* 0111:B4; Sigma–Aldrich), which is a Toll-like receptor (TLR)4 agonist used as a positive control for fully matured cells.

Plaque Assay

Vero cells cultured on 24-well plates were treated with 10-fold serial dilutions of ECTV stocks obtained from C57BL/6 and BALB/c GM-BM at 4, 12, and 24 hpi. After 5 days, plaques were counted under Olympus IX71 inverted microscope. After counting, Vero cell monolayers were stained with 0.3% crystal violet and air dried.

Measurement of Apoptosis

The apoptotic rate of GM-BM was measured using the FITC Annexin V Apoptosis Detection Kit I (BD Biosciences),

according to the manufacturer's protocol. Briefly, cells were washed twice with cold phosphate-buffered saline (Sigma–Aldrich) and resuspended in 1× binding buffer at a concentration of 1×10^6 cells/ml. Then, 100 μl of the solution (containing 1×10^5 cells) was stained with 5 μl Annexin V-FITC and 5 μl propidium iodide (PI) and incubated for 15 min at room temperature in the dark. Finally, cells were resuspended in 400 μl of 1× binding buffer and analyzed immediately by flow cytometry. Viable cells (no measurable apoptosis) are Annexin V-FITC and PI negative, early apoptotic cells (membrane integrity is present) are Annexin V-FITC positive and PI negative, whereas late apoptotic cells (end stage apoptosis and death) are Annexin V-FITC and PI positive.

Real-Time Reverse Transcription Polymerase Chain Reaction (RT-PCR)

The expression of selected genes involved in innate and adaptive immune functions of GM-BM was evaluated using real-time reverse transcription polymerase chain reaction (RT-PCR), as previously described (Dolega et al., 2017; Szulc-Dąbrowska et al., 2017) with minor modifications. Briefly, RNA was extracted from 1×10^6 mock- or ECTV-infected GM-BM, untreated or treated with LPS for 24 h using Qiagen RNeasy Mini Kit (Qiagen, Inc., Valencia, CA, United States), as recommended by the manufacturer. Additionally, on-column DNase I digestion of genomic DNA was performed using RNase-Free DNase Set (Qiagen). The RNA concentration and purity was assessed using the Take-3 system on Epoch BioTek spectrophotometer and analyzed in Gen5 software (BioTek Instruments, Inc., Winooski, VT, United States). Total RNA (1 μg) was applied for cDNA synthesis using the RT^2 First Strand Kit (Qiagen) according to the manufacturer's protocol. Before RT, genomic DNA was additionally removed by incubation in GE2 buffer for 5 min at 42°C. For analysis of selected genes involved in regulation of innate and adaptive immune properties of GM-BM the Mouse Dendritic and Antigen Presenting RT^2 Profiler PCR Array (Qiagen) was used according to the recommendation of the manufacturer. Briefly, 550 ng cDNA was mixed with RT^2 SYBR Green Mastermix (Qiagen) and aliquoted into the 96-well RT^2 Profiler PCR array plate, containing lyophilized RT^2 qPCR primers for a set of 84 related genes, five housekeeping genes (*Actb*, *B2m*, *Gapdh*, *Gusb*, and *Hsp90ab1*), three Reverse Transcription Controls (RTC), three Positive PCR Controls (PPC), and one Mouse Genomic DNA Contamination (MGDC) control (**Supplementary Table S1**). Amplification was performed in ABI 7500 thermocycler (Life Technologies, Carlsbad, CA, United States) at 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 1 min. Fluorescence data were collected each cycle after the 1 min step at 60°C. Amplification data were acquired through SDS Software (Applied Biosystems).

Data Quality Control, Normalization, and Analysis

Obtained data fulfilled the criteria of PCR array reproducibility, RT efficiency, and genomic DNA contamination. If the average PPC Ct was 19 ± 3 and no two arrays had an average PPC Ct > 2 away from one another, then the samples passed the criteria for the PCR array reproducibility. If ΔCt (AVG RTC − AVG PPC) was ≤5, then the samples passed the criterion for RT efficiency. If Ct (MGDC) was ≥35, then the samples passed the criterion for genomic DNA contamination.

Three independent biological experiments were performed for each experimental group. The normalization was performed using the most stable genes/gene in the PCR array data set, identified by software at the Qiagen Data Analysis Center. The Ct values for these genes were geometrically averaged and used for the calculation of ΔΔCt values. The data are presented as fold change ($2^{-\Delta\Delta Ct}$) which is the normalized gene expression ($2^{-\Delta Ct}$) in the test sample divided by the normalized gene expression ($2^{-\Delta Ct}$) in the control sample. Fold regulation represents fold change results in a biologically meaningful way. Fold change values greater than one indicate a positive- or an up-regulation, and the fold regulation is equal to the fold change. Fold change values less than one indicate a negative or down-regulation, and the fold regulation is the negative inverse of the fold-change. The *P*-values were calculated based on a Student's *t*-test of the replicate $2^{-\Delta Ct}$ values for each gene in the control group and treatment groups. Significance was assessed at $^*P \leq 0.05$ and $^{**}P \leq 0.01$.

Immunophenotyping

Multi-color immunophenotyping was performed as previously described (Szulc-Dąbrowska et al., 2017). Cell surface maturation markers were stained with the following mAbs used in appropriate combinations: anti-H-2D[b]-PE (KH95, mouse IgG2b), anti-H-2D[d]-PE (34-2-12, mouse IgG2a), anti-I-A/I-E-BV711, anti-CD40-APC (3/23, rat IgG2a), anti-CD80-APC (16-10A1, Armenian hamster IgG2), anti-CD83-PE (Michel 19, rat IgG1) (all from BD Biosciences), and anti-CD86-PerCP-Cy5.5 (GL-1, rat IgG2a; eBioscience). Chemokine receptors were stained with anti-CD191(CCR1)-APC (643854, Rat IgG2b; R&D Systems), anti-CD195(CCR5)-PE (C34-3448, rat IgG2c; BD Biosciences), and anti-CD197(CCR7)-PerCP/Cy.5.5 (4B12, rat IgG2a; BioLegend) mAbs. Additionally, GM-BM were labeled for CD11c marker using anti-CD11c-BV421 mAbs. Appropriate isotype controls (purchased from BD Biosciences) were used to determine the percentages of cells expressing the respective markers. Moreover, to obtain proper gating strategies the Fluorescence Minus One samples were included as negative controls.

Intracellular Staining

Non-, uvi-ECTV-, or ECTV-infected GM-BM were left unstimulated or were stimulated with LPS for 24 h in a 24-well plate at a density of 5×10^5 cells/well. Brefeldin A (6 μg/ml; BD Biosciences) was added for the last 5 h of culture. Cells were then collected and stained with anti-CD11c-BV421. Intracellular staining of cytokines was performed using a Cytofix/Cytoperm kit (BD Biosciences) according to the manufacturer's instructions. The following mAbs were used for cytokine detection: anti-tumor necrosis factor (TNF)-FITC (MP6-XT22, rat IgG1), anti-IL-12(p40/p70)-APC (C15.6, rat IgG1; both from BD Biosciences) and anti-CCL3 [macrophage

inflammatory protein 1 alpha (MIP-1α)]-PE (DNT3CC; rat IgG2a; eBioscience, San Diego, CA, United States). In some experiments, cells were stained intracellularly for the presence of ECTV antigens using a Cytofix/Cytoperm kit and rabbit polyclonal Abs anti-ECTV-FITC, obtained as previously described (Szulc-Dabrowska et al., 2016). The staining procedure also included appropriate isotype controls obtained from BD Biosciences and eBioscience.

Flow Cytometry Analysis

The population of live cells was gated based on size and granularity according to forward (FSC-A) and side (SSC-A) scatter profile. GM-BM enriched in cDC population were gated as FSC^{high} and $CD11c^{+}$. Twenty thousand gated cell events were acquired from each specimen and analyzed on a BD FACSCanto II and BD LSRFortessa flow cytometers (Becton Dickinson). Data were analyzed with FACSDiva 7.0 software (Becton Dickinson).

Enzyme-Linked Immunosorbent Assay

$CD11c^{+}$-enriched GM-BM were plated into a 24-well plate at a density of 2×10^4 cells/well. After treatment with medium (mock), uvi-ECTV, or live ECTV, cells were incubated with or without LPS for 4, 8, 12, 18, and 24 h. In some experiments, cells were additionally stimulated with 3 mM ATP for 6 h to induce IL-1β and IL-18 release (Mehta et al., 2000). The concentration of cytokines and chemokines in culture supernatants was quantified using available commercial sandwich enzyme-linked immunosorbent assay (ELISA), according to the instructions given by the manufacturers. TNF, IL-6, IL-10, IL-12p40, IL-12p70, and CCL2/monocyte *chemoattractant* protein 1 (MCP-1) were detected using BD OptEIA ELISA sets (BD Biosciences). CCL3/MIP-1α and CCL5/RANTES (regulated upon activation normal T cell expressed and secreted) were quantified using Quantikine ELISA Kits (R&D Systems). IL-18 and IL-15R/IL-15 complex were assessed using Platinum ELISA and ELISA Ready-SET-Go!, respectively (eBioscience). The absorbance was read at 450 nm using Epoch Microplate Spectrophotometer. Each plate contained its own standard curve for determination of cytokine/chemokine levels in the supernatants.

Nitric Oxide detection

Nitric oxide (NO) formation can be investigated by measuring nitrite (NO_2^-), which is one of two primary, stable, and non-volatile breakdown products of NO. The amount of nitrite (μM/ml) in culture supernatants was assessed using the Griess Reagent System (Promega, Madison, WI, United States), according to the manufacturers' protocol. Briefly, an equal volume of supernatant was mixed with the Sulfanilamide Solution and incubated for 10 min at room temperature in the dark. Then, *N*-1-naphthylethylenediamine dihydrochloride solution was added and incubated for additional 10 min. The absorbance was measured at 520 nm using Epoch Microplate Spectrophotometer. A standard curve generated with sodium nitrite was used to calculate the nitrite concentration in the supernatants.

Statistical Analysis

Data are presented as mean ± standard deviation (SD) from at least three independent biological replicates. Normal distribution of variables was assessed using the Shapiro–Wilk *W*-test. If the data were normally distributed and the variances were homogeneous, then two-independent (unpaired) Student's *t*-test was applied for group comparison. Some data were analyzed using two-dependent (paired) Student's *t*-test. If the data were non-normally distributed, then the Wilcoxon signed-rank or Mann–Whitney *U*-tests were applied for comparison of paired and unpaired values, respectively (STATISTICA 6.0 software, StatSoft Inc., Tulsa, OK, United States). Statistical significance was assessed at $^{*}P \leq 0.05$ and $^{**}P \leq 0.01$.

RESULTS

The Kinetics of ECTV Replication Cycle Are Similar in GM-BM Derived from C57BL/6 and BALB/c Mice

Our previous study showed that ECTV is able to productively infect murine GM-BM, including cDCs, with subsequent release of progeny virions (Szulc-Dąbrowska et al., 2017). In the present study, we compared the kinetics of ECTV replication cycle in GM-BM derived from resistant C57BL/6 and susceptible BALB/c mice. Cells were infected with ECTV at MOI = 1 and after 4, 12, and 24 hpi the percentage of $ECTV^{+}$ cells was determined by intracellular staining and flow cytometry analysis (**Figure 1A**). The mean percentage of $ECTV^{+}$ cells during the first 12 h of virus replication was comparable between both strains of mice (**Figure 1B**). At 24 hpi, the percentage of $ECTV^{+}$ cells was slightly higher in C57BL/6 comparing to BALB/c mice (75 vs. 63%, respectively). LPS treatment of ECTV-infected GM-BM derived from both mouse strains significantly ($P \leq 0.05$) increased the percentage of $ECTV^{+}$ cells during the entire virus replication cycle. It suggests that TLR4-agonist stimulation positively regulates the tempo of virus reproduction in GM-BM infected at MOI of 1 (Szulc-Dąbrowska et al., 2017). Additionally, to confirm that ECTV infection of GM-BM is productive, we performed a plaque assay to determine quantify infectious viral particles (**Figure 1C**). At 4 hpi, the number of infectious virions was 4.2×10^3 and 5×10^3 PFU/ml in C57BL/6 and BALB/c GM-BM cultures, respectively. At 12 and 24 hpi, the virus titers in both cultures increased around 100- and 500-fold, respectively, and reached 5×10^5 and 7×10^5 PFU/ml at 12 hpi, and 3.2×10^6 and 2.6×10^6 PFU/ml at 24 hpi in C57BL/6 and BALB/c GM-BM cultures, respectively (**Figure 1C**). LPS treatment slightly increased the number of infectious virus particles, especially at 24 hpi and is agreement with flow cytometry results.

Our previous study also showed that 10-day culture of GM-BM and later infected with ECTV at MOI = 5 exhibited a high percentage (more than 30%) of early and late apoptotic cells during later stages of infection (Szulc-Dąbrowska et al., 2017). To minimize the apoptotic rate in GM-BM cultures, in the present study we used 8-day culture of GM-BM and later infected with

ECTV at MOI = 1. Under such conditions, we observed only a small induction of apoptosis in GM-BM at 24 hpi in the absence of LPS (**Figure 1D**). Meanwhile, ECTV infection of GM-BM in the presence of LPS slightly (but not significantly) increased the percentage of apoptotic cells compared with LPS-treated mock- or uvi-ECTV-infected cells (**Figure 1D**). It is likely that LPS prevents the apoptotic effect induced by the virus infection, since LPS has been shown to inhibit caspase 3-dependent apoptosis in immune cells (Russe et al., 2014).

Generalized Down-Regulation of Genes Involved in Innate and Adaptive Immune Functions Is Comparable in GM-BM-Derived from C57BL/6 and BALB/c Mice

Before assessment of ECTV impact on gene expression profile, we first compared the gene expression between uninfected GM-BM derived from C57BL/6 and BALB/c mice to determine interstrain differences in genes involved in innate and adaptive immune functions (**Figure 2**). These genes, encoding proteins that regulate biological properties of DCs, were divided into seven categories: (1) antigen uptake, (2) antigen presentation, (3) chemotaxis, (4) chemokines and cytokines, (5) cytokine receptors, (6) other cell surface receptors, and (7) signal transduction.

Before calculating $2^{-\Delta\Delta Ct}$ values, the data were normalized to the geometric mean of the four most stable housekeeping genes: *ActB*, *Gapdh*, *Gusb*, and *Hsp90ab1*. The scatter plots of the mRNA expression profiles of untreated and LPS-treated C57BL/6 vs. BALB/c GM-BM are presented in **Figures 2A,D**, respectively. The scatter plots compare the normalized expression of every gene on the array between two groups by plotting them against one another to quickly visualize large gene expression changes. As shown in **Figure 2A**, C57BL/6 compared to BALB/c GM-BM exhibited up-regulation of 25 genes, including *Ccl5*, *Ccl8*, *Ccl17*, *Cxcl10*, *Cd40*, *Fas*, *Fcer2a*, *Fcgr1*, *Il-12b*, *Irf7*, *Thbs1*, and *Tlr9*, and down-regulation of four genes, *Cd1d2*, *Cd2*, *Fcgrt*, and *Icam2*. LPS treatment of C57BL/6 and BALB/c GM-BM for 24 h resulted in up-regulation of 42 and 41 genes, and down-regulation of 13 and 15 genes, respectively (**Figures 2B,C**). Up-regulated genes were those known to influence maturation of DCs and included genes for cell surface receptors (e.g., *Cd1d1*, *Cd1d2*, *Cd40*, *Cd80*, *Cd86*, *H2-DMa*, *Tlr1*, *Tlr2*), chemokines and cytokines (e.g., *Ifn*, *Il10*, *Il12a*, *Il12b*, *Tnf*, *Ccl2*, *Ccl3*, *Ccl4*, *Ccl5*, *Ccl7*, *Ccl8*, *Ccl12*, *Ccl17*, *Ccl17*, *Cxcl1*, *Cxcl2*, *Cxcl10*) and signal transduction (e.g., *Irf7*). *Cd28*, *Cd36*, *Cd209a*, and *Clec4b2* were the commonly down-regulated genes after LPS exposure of GM-BM from both mouse strains. Interestingly, C57BL/6 GM-BM treated with LPS for 24 h showed up- and down-regulation of 31 and 5 genes, respectively, compared to BALB/c GM-BM treated with LPS (**Figure 2D**). Among up-regulated genes there were those primarily engaged in the antigen uptake (*Tap2*) and presentation (*Cd4*, *Cd40*, *Cd80*, *Cd86*, *H2-DMa*, *Thbs1*) and chemokine and cytokine production (*Ccl5*, *Ccl7*, *Ccl8*, *Ccl12*, *Cxcl2*, *Cxcl10*, *Cxcl12*). Moreover, TLR9 mRNA was significantly ($P \leq 0.01$) up-regulated in C57BL/6 compared with BALB/c GM-BM after LPS-treatment. Collectively, it may suggest that C57BL/6 GM-BM are at a higher maturation state than BALB/c GM-BM after TLR4 agonist stimulation.

The gene expression comparison between uninfected and infected cells untreated or treated with LPS was made after data normalization to *Cdkn1a*, due to its uniform expression across treatment groups in the experiments. Differentially expressed genes are presented by volcano plots, which summarize fold-change in gene expression and statistical significance in four different experimental groups (**Figure 3**). ECTV infection of GM-BM derived from C57BL/6 and BALB/c mice resulted in a global down-regulation of 75 and 74 genes, respectively, among which 52 and 50 were significantly ($P \leq 0.05$) repressed (**Figures 3A,B**). Profound inhibition in gene expression was observed in all gene categories in cells from both mouse strains. Only *Il10* gene was significantly ($P \leq 0.05$) up-regulated in infected cells. Moreover, ECTV infection inhibited LPS-induced up-regulation of mRNA transcripts for many genes involved in activation and maturation of GM-BM. LPS-treated ECTV-infected cells derived from C57BL/6 and BALB/c mice displayed down-regulation of 65 and 63 genes, respectively, compared to LPS-treated uninfected cells (**Figures 3C,D**). Among these genes 45 and 48, respectively, were significantly ($P \leq 0.05$) down-regulated. Taken together, our data clearly demonstrate that ECTV induces global gene repression in GM-BM of both strains and this repression is maintained during TLR4 agonist treatment.

Expression of Genes Involved in Antigen Uptake and Presentation

In the antigen uptake category, *Cd44*, *Icam1*, *Icam2*, and *Rac1* were mostly significantly ($P \leq 0.05$) down-regulated in untreated or LPS-treated ECTV-infected C57BL/6 and BALB/c GM-BM in comparison with their uninfected counterparts (**Figure 4A**). The mean expression level of these genes was more than 10-fold repressed. The level of *Rac1* repression was significantly ($P \leq 0.05$) higher in infected C57BL/6 than BALB/c cells. Meanwhile, *Cdc42* was not significantly up-regulated in LPS-treated ECTV-infected GM-BM from C57BL/6 mice compared to LPS-treated uninfected cells (**Figure 4A**).

Of the 16 genes analyzed in the antigen presentation category, ECTV infection resulted in significantly ($P \leq 0.05$) down-regulation of 11 mRNA transcripts in C57BL/6 and BALB/c GM-BM (**Figure 4B**). *B2m*, *Cd1d1*, *Cd74*, *Cd80*, *Cd86*, *Fcgrt*, *H2-DMa*, *Tapbp*, and *Thbs1* were significantly ($P \leq 0.05$) repressed in cells from both mouse strains, whereas *Cd28*, *Cd40Ig*, and *Cd1d2*, *Cd40* were significantly down-regulated in C57BL/6 and BALB/c GM-BM, respectively. The level of *Cd74* repression was significantly ($P \leq 0.05$) higher in infected BALB/c than C57BL/6 cells. LPS treatment of mock-infected C57BL/6 and BALB/c GM-BM resulted in up-regulation of the following genes: *Cd1d1*, *Cd1d2*, *Cd209a*, *Cd40*, *Cd80*, *Cd86*, *Tapbp*, and *Thbs1*. Meanwhile, treatment of ECTV-infected cells with LPS did not up-regulate genes, as observed in mock-infected LPS-treated GM-BM from the two mouse strains. After stimulation with LPS, infected C57BL/6 and BALB/c GM-BM still exhibited a profound down-regulation of 10 following genes: *B2m*, *Cd1d1*,

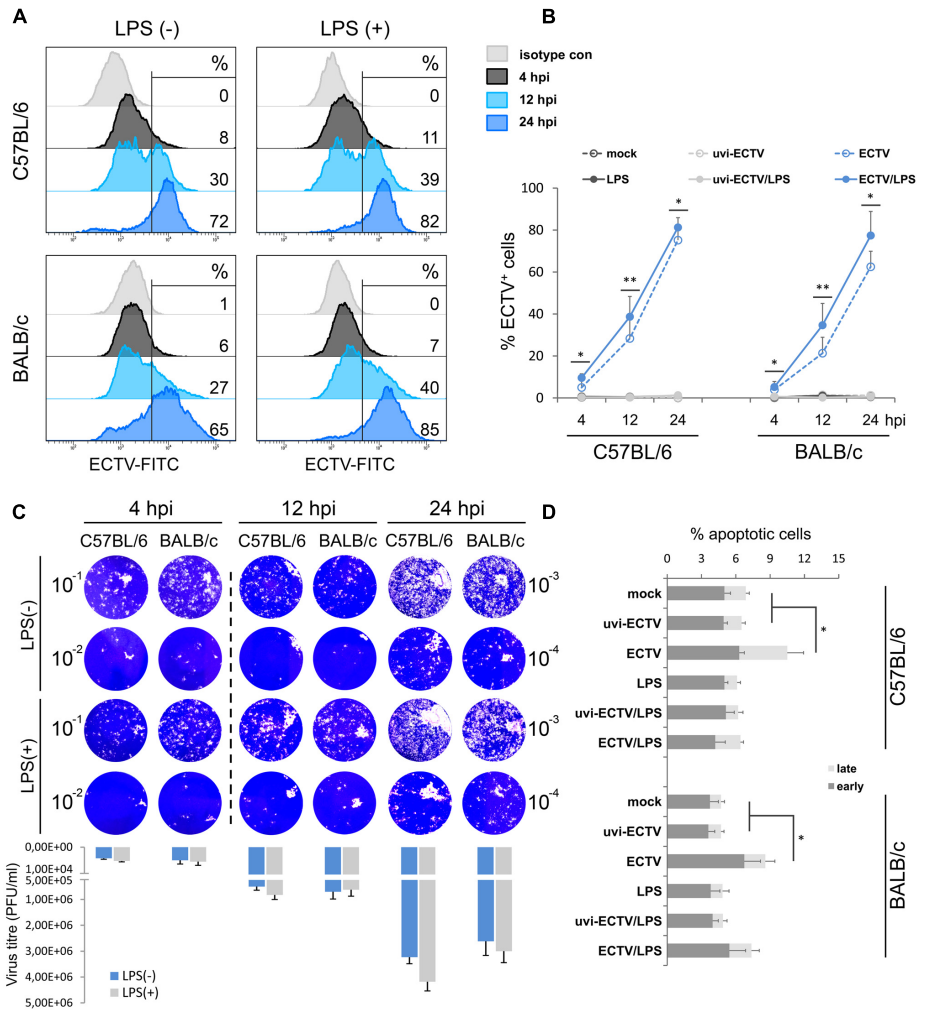


FIGURE 1 | Similar kinetics of ECTV replication in GM-BM of C57BL/6 and BALB/c mice. **(A)** Representative flow cytometry histograms showing the percentage of ECTV$^+$ cells at 4, 12, and 24 hpi of GM-BM. **(B)** The percentage (mean ± SD) of ECTV$^+$ cells in GM-BM culture. **(C)** Plaque assay determination of ECTV titer in GM-BM at 4, 12, and 24 hpi. Plaques were visualized by staining of Vero cell monolayers with 0.3% crystal violet. The mean number of infectious virions (PFU/ml) was calculated from two experiments. **(D)** The percentage (mean ± SD) of early and late apoptotic cells in GM-BM during ECTV infection (Student's *t*-test; $^*P < 0.05$, $^{**}P < 0.01$).

Cd1d2, *CD40*, *Cd74*, *Cd86*, *Fcgrt*, *H2-DMa*, *Tapbp*, and *Thbs1*. In all cases the most repressed genes were *Cd1d1* and *Thbs1*, and their mean expression levels were ≥300-fold decreased. The repression level of mRNA transcript for CD40 was significantly ($P \leq 0.05$) higher in LPS-stimulated ECTV-infected C57BL/6 than BALB/c cells. The mRNA expression of CD80 was not significantly ($P > 0.05$) down-regulated (**Figure 4B**). Only *Cd4* was differentially regulated in LPS-treated ECTV-infected GM-BM from the two strains of mice. *Cd4* was significantly ($P = 0.009$) up-regulated in BALB/c cells, whereas in C57BL/6 cells its expression remained unchanged.

Expression of Genes Involved in DC Chemotaxis

The four important genes in this group were significantly ($P \leq 0.05$) down-regulated in ECTV-infected GM-BM from both mouse strains (**Figure 4C**). These genes were *Ccl5*, *Ccr1*, *Ccr2*, and *Ccr5*. *Ccr1* was significantly ($P \leq 0.05$) more repressed in infected C57BL/6 than BALB/c cells. The same set of genes were also down-regulated in LPS-treated ECTV-infected GM-BM compared to LPS-treated uninfected cells. However, in cells from C57BL/6 mice, the repression of *Ccr2* and *Ccr5* was not statistically significant ($P > 0.05$). Although the down-regulation of *Cxcr4* in GM-BM from both mouse strains was not significant, statistical analysis revealed that this gene was considerably repressed in C57BL/6 mice.

Expression of Genes Involved in Chemokine and Cytokine Production

The next category of analyzed genes concerned those engaged in the synthesis of chemokines (**Figure 5A**) and cytokines (**Figure 5B**). Among the 15 chemokine genes assayed, seven

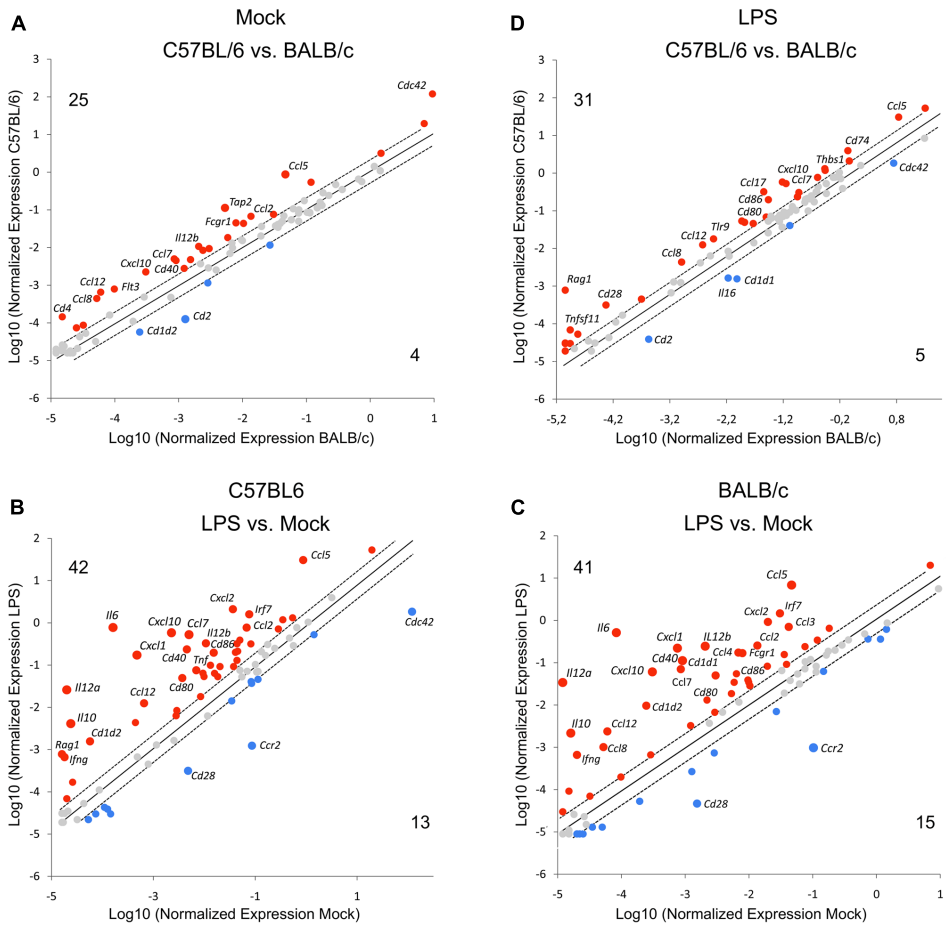


FIGURE 2 | Differential expression of genes involved in innate and adaptive immune functions in GM-BM from C57BL/6 and BALB/c mice. The scatter plots of gene expression profiling in **(A)** mock-infected GM-BM from C57BL/6 vs. BALB/c mice, **(B)** LPS-treated vs. untreated mock-infected GM-BM from C57BL/6 mice, **(C)** LPS-treated vs. untreated mock-infected GM-BM from BALB/c mice, **(D)** mock-infected LPS-treated (for 24 h) GM-BM from C57BL/6 vs. BALB/c mice. The numbers represent genes that are up-regulated (red) and down-regulated (blue). The cut-offs for twofold induction and repression are indicated by dotted lines (all points having a fold-change less than 2 are shown in gray).

genes in both C57BL/6 and BALB/c GM-BM were significantly ($P \leq 0.05$) down-regulated (**Figure 5A**) following ECTV infection. *Ccl4*, *Ccl5*, *Ccl17*, and *Ccl20* were repressed in cells from both mouse strains, whereas *Ccl2*, *Ccl3*, *Ccl8* and *Ccl12*, *Cxcl2*, *Cxcl12* were decreased in C57BL/6 and BALB/c GM-BM, respectively (**Figure 5A**). LPS treatment of mock-infected cells up-regulated 12 genes (*Ccl2*, *Ccl3*, *Ccl4*, *Ccl5*, *Ccl7*, *Ccl8*, *Ccl12*, *Ccl17*, *Ccl20*, *Cxcl1*, *Cxcl2*, and *Cxcl10*) in GM-BM from both mouse strains (**Figures 2C,D**). Meanwhile, LPS-treated ECTV-infected cells showed significant ($P \leq 0.05$) down-regulation of seven mRNA transcripts compared to LPS-treated uninfected cells (**Figure 5A**). GM-BM from both mouse strains exhibited repression of *Ccl2*, *Ccl3*, *Ccl5*, and *Ccl8*, whereas *Ccl12*, *Cxcl1*, *Cxcl10* and *Ccl4*, *Ccl17*, *Cxcl2* were significantly ($P \leq 0.05$) decreased in cells from C57BL/6 and BALB/c mice, respectively. ECTV-infected C57BL/6 GM-BM displayed significant ($P \leq 0.05$) down-regulation of *Ccl2*, *Cxcl12* and *Ccl2*, *Ccl3*, *Ccl8* than BALB/c cells, cultured without or with LPS, respectively.

Among the 13 genes analyzed in the cytokine group, *Ifng*, *Mif*, *Tgfb1*, and *Tnf* were significantly ($P \leq 0.05$) repressed in GM-BM of both mouse strains (**Figure 5B**). The level of *Mif* expression was significantly ($P \leq 0.05$) decreased in C57BL/6 cells. *Il12a*, *Il16*, and *Flt3l* were also down-regulated, however, significant ($P \leq 0.05$) repression of these genes was observed only in C57BL/6 cells. On the contrary, the expression of *Il12b* was significantly ($P \leq 0.01$) decreased in BALB/c GM-BM. Interestingly, ECTV infection significantly ($P \leq 0.05$) up-regulated the expression of *Il10* in cells from both mouse strains (**Figure 5B**).

Uninfected GM-BM from C57BL/6 and BALB/c mice displayed up-regulation of *Ifng*, *Il6*, *Il10*, *Il12a*, *Il12b*, and *Tnf* after LPS treatment for 24 h (**Figures 2C,D**). Meantime, LPS-stimulated ECTV-infected cells exhibited significant ($P \leq 0.05$) repression of *Il6*, *Il10*, *Il12a*, *Mif*, and *Tgfb1*, compared to uninfected cells incubated with TLR4 agonist (**Figure 5B**). Moreover, LPS-treated ECTV-infected GM-BM from C57BL/6 and BALB/c mice showed down-regulation of *Il16* and *Flt3l* or

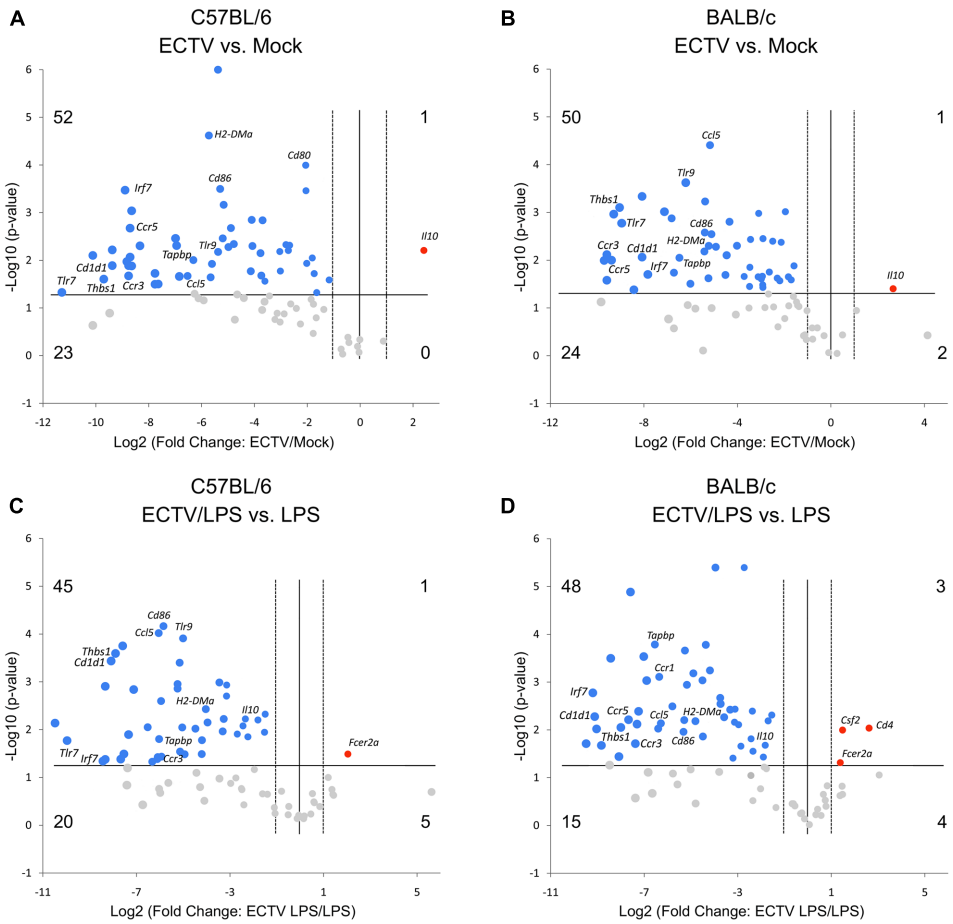


FIGURE 3 | ECTV similarly affects multiple genes involved in innate and adaptive immune functions in GM-BM from C57BL/6 and BALB/c mice. The volcano plots show differentially expressed genes in ECTV- vs. mock-infected GM-BM from C57BL/6 **(A)** and BALB/c **(B)** mice, and ECTV + LPS- vs. LPS-treated GM-BM from C57BL/6 **(C)** and BALB/c **(D)** mice. The volcano plots display the threshold for statistical significance (log10 of the *P*-value; horizontal line) and fold change (log2; vertical dotted lines). The two vertical dotted lines represent fold changes of –2 and +2, the horizontal line represents significant *P*-value of 0.05. The number of genes that are down(blue)- and up(red)-regulated by at least twofold and have a *P*-value less than 0.05 are at the upper-left and upper-right, respectively. The number of genes that are down- and up-regulated (gray) by at least twofold, however, have a *P*-value more than 0.05 are at the lower-left and lower-right, respectively.

Ifng and *Tnf*, respectively. *Csf2* was significantly up-regulated only in BALB/c GM-BM after ECTV infection and LPS treatment.

Expression of Genes for Cytokine and Other Cell Surface Receptors

In the cytokine receptors category, *Ccr1*, *Ccr2*, *Ccr3*, *Ccr5*, *Csf1r*, and *Lyn* were significantly ($P \leq 0.05$) down-regulated in untreated or LPS-treated ECTV-infected GM-BM from both mouse strains, compared to their uninfected counterparts (**Figure 6A**). *Ccr1*, *Cxcr4*, and *Lyn* were significantly ($P \leq 0.05$) more repressed in C57BL/6 than BALB/c GM-BM. Moreover, *Ccr9* and *Flt3* were significantly decreased only in infected BALB/c cells.

The majority of other analyzed genes in the cell surface receptors category were down-regulated in ECTV-infected GM-BM from both mouse strains (**Figure 6B**). *Cd36*, *Fcgr1*, *Lrp1*, *Tlr1*, and *Tlr2* were significantly ($P \leq 0.05$) repressed in cells from both mouse strains and such inhibitory effect was independent of LPS stimulation. *Cd2*, *Tlr7*, and *TLr9* were also down-regulated in both strains, however, after LPS stimulation *Tlr7* and *TLr9* were significantly ($P \leq 0.05$) repressed only in C57BL/6 GM-BM. On the contrary, *Fcer2a* was significantly ($P \leq 0.05$) up-regulated in ECTV-infected C57BL/6 and BALB/c GM-BM treated with LPS (**Figure 6B**). LPS stimulation of mock-infected cells resulted in up-regulation of *Cd40*, *Fcgr1*, *Tlr1*, and *Tlr2* (**Figures 2C,D**).

Expression of Genes Involved in Signal Transduction

The last analyzed group concerned 12 genes involved in signal transduction (**Figure 6C**). The following 8 genes were mostly significantly ($P > 0.05$) down-regulated in ECTV-infected GM-BM derived from both strains of mice and independent of LPS treatment: *Fas*, *Irf7*, *Itgam*, *Itgb2*, *Nfkb1*, *Ptprc*, *Rela*, and *Stat3*. Moreover, *Cebpa* and *Relb* were significantly ($P > 0.05$) down-regulated in GM-BM from both and only BALB/c mouse,

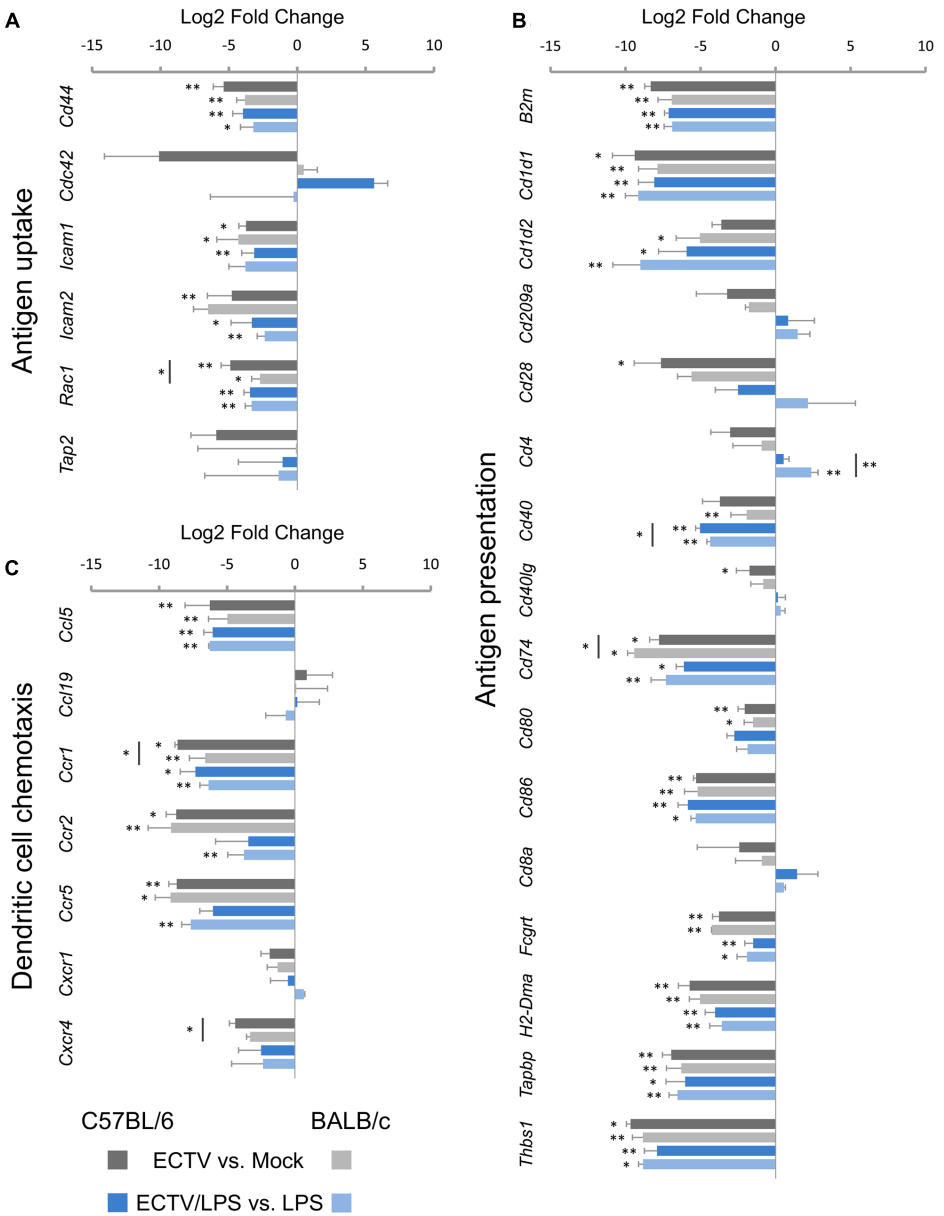


FIGURE 4 | Mouse strain-independent effect of ECTV infection on the expression of genes involved in **(A)** antigen uptake, **(B)** antigen presentation, and **(C)** chemotaxis in GM-BM (unpaired Student's *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$).

respectively. *Cebpa*, *Irf7* and *Stat3* were significantly ($P > 0.05$) more repressed in infected C57BL/6 than BALB/c GM-BM (**Figure 6C**).

Mouse Strain-Independent Effect of ECTV-Infection on Cytokine and NO Production by GM-BM

cDCs play a key role in driving T cell responses by eliciting polarizing signals, the most important of which are cytokines that selectively promote the generation of Th1 or Th2 cells. We next confronted the GM-BM obtained from resistant C57BL/6 and susceptible BALB/c mice regarding their Th-polarizing cytokine profile under ECTV infection *in vitro*. The analyzed cytokines and chemokines are engaged in T cell activation and involved in regulation of Th1 or Th2 polarization during the adaptive immune responses.

TNF-α contributes to DC activation and maturation and is required for subsequent induction of optimal T cell responses (Zhang et al., 2003). Uninfected GM-BM from C57BL/6 and BALB/c mice produced TNF-α at similar low levels, however, after LPS stimulation, BALB/c cells secreted higher amounts of TNF-α than C57BL/6 cells (**Figure 7A**). During LPS treatment, the concentration of TNF-α in culture

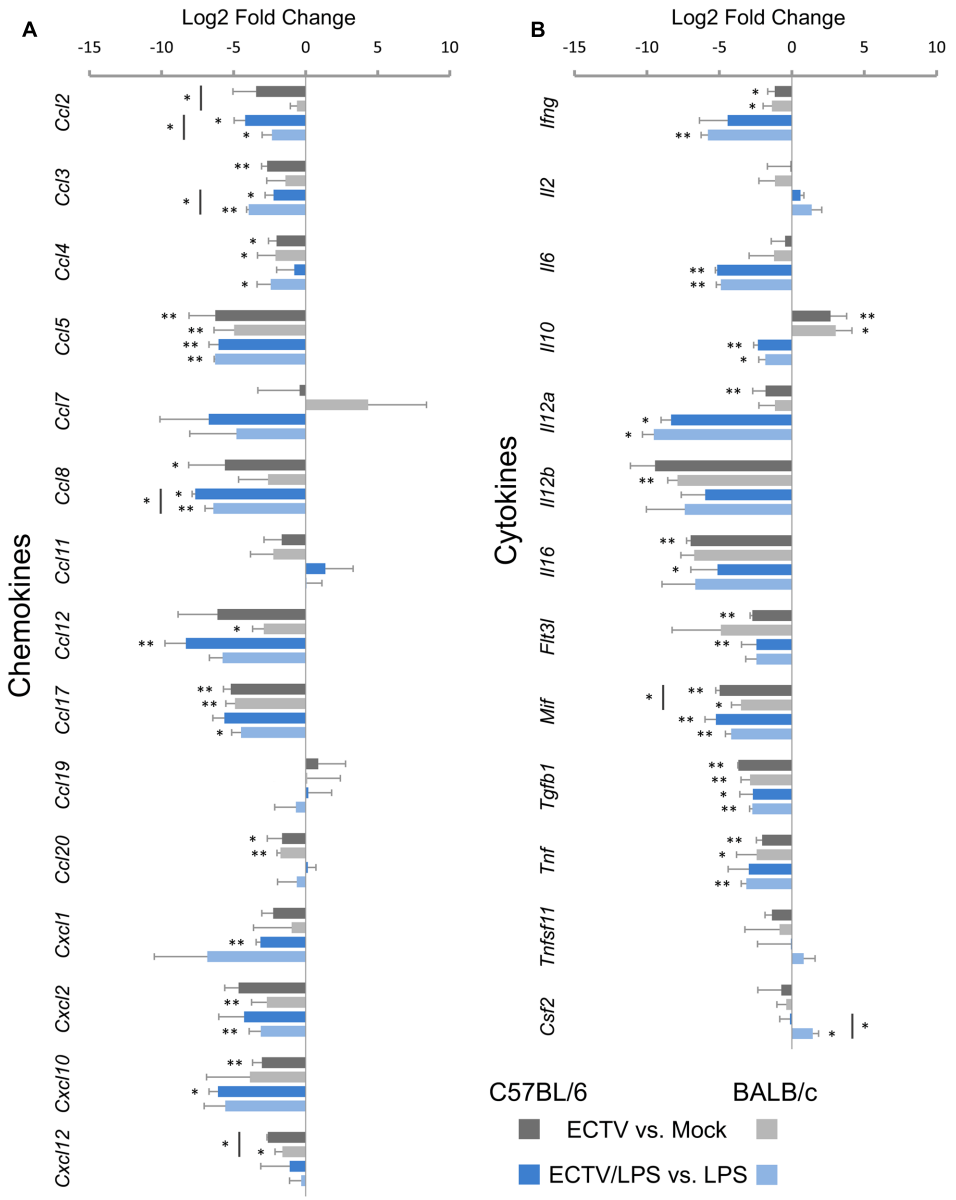


FIGURE 5 | Mouse strain-independent effect of ECTV infection on the expression of genes for chemokines **(A)** and cytokines **(B)** in GM-BM (unpaired Student's *t*-test; $*P < 0.05$, $**P < 0.01$).

supernatants from C57BL/6 and BALB/c GM-BM started to decrease at 24 h post-stimulation and was two- and fourfold lower, respectively, than at 4 h post-stimulation. Meantime, ECTV infection of GM-BM from both strains of mice profoundly inhibited TNF-α secretion, even after LPS treatment.

IL-6 produced by DCs in response to TLR recognition of microbial products has been demonstrated to regulate T cell activation by overcoming the suppressive effect of Tregs (Pasare and Medzhitov, 2003). Similar to TNF-α, mock- or uvi-ECTV-infected GM-BM from both mouse strains secreted comparable low levels of IL-6. After LPS stimulation, GM-BM of both strains of mice produced high amounts of IL-6 and from 8 h post-stimulation, DCs from BALB/c mice produced higher levels of IL-6 than those from C57BL/6 mice (**Figure 7A**). ECTV-infected GM-BM untreated or treated with LPS were able to secrete IL-6, however, the amounts were significantly ($P > 0.01$) lower compared to those secreted by mock- or uvi-ECTV-infected LPS-untreated or -treated GM-BM.

Next, we analyzed cytokines and chemokines engaged in Th1 polarization, such as: IL-12p40, IL-12p70, IL-15/IL-15Rα, IL-18, CCL3/MIP-1α, and CCL5/*RANTES* (**Figures 7A,B**). C57BL/6 GM-BM produced larger amounts of IL-12p40 and IL-12p70 than BALB/c GM-BM, especially after LPS treatment

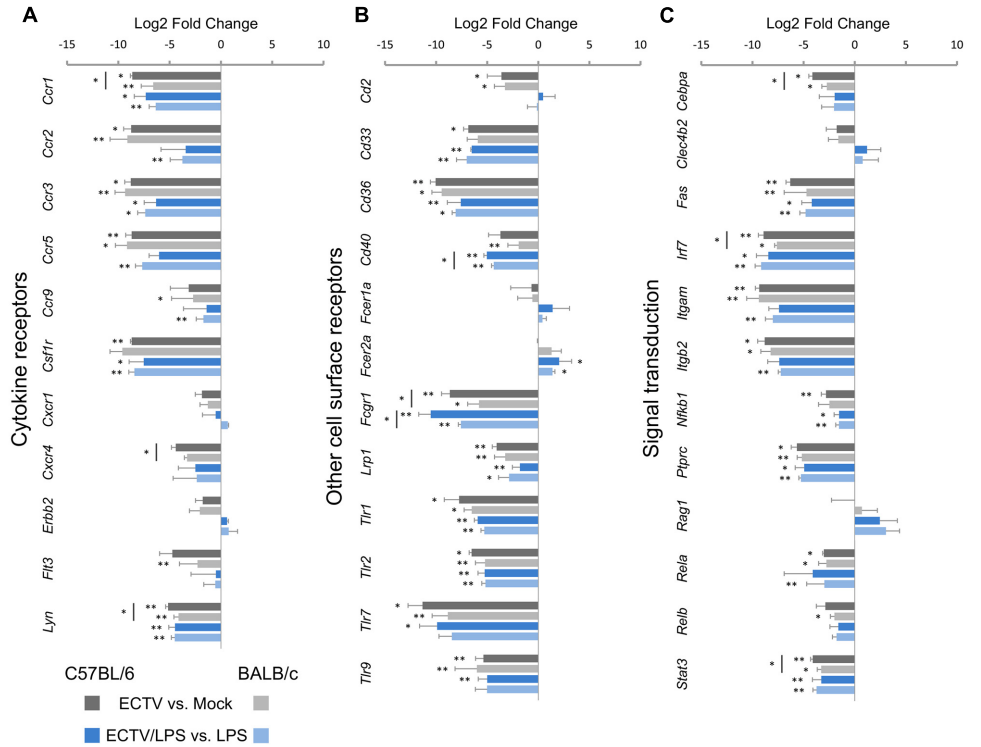


FIGURE 6 | ECTV similarly affects genes for **(A)** cytokine receptors and **(B)** other cell surface receptors, and **(C)** genes involved in signal transduction in GM-BM from C57BL/6 and BALB/c mice (unpaired Student's *t*-test; $*P < 0.05$, $**P < 0.01$).

(**Figure 7A**). ECTV-infection significantly ($P > 0.01$) decreased the production of both cytokines by LPS-untreated or -treated GM-BM of both mouse strains. Meantime, CCL3 was secreted at higher levels by uninfected BALB/c than C57BL/6 cells, especially in response to the ligand for TLR4 (**Figure 7A**). At 12 hpi with ECTV, GM-BM from both mouse strains produced less CCL3 than mock- or uvi-ECTV infected cells, whereas at 24 hpi those cells showed significantly ($P > 0.01$) increased level of CCL3 compared to mock- or uvi-ECTV-treated cells. However, GM-BM infected with live ECTV exhibited the suppression of TLR4 agonist-induced secretion of CCL3. CCL5 production was higher by C57BL/6 GM-BM, however, LPS stimulation induced robust secretion of comparable large amounts of CCL5 by cells from both mouse strains (**Figure 7A**). The secretion of CCL5 was >2-fold reduced in ECTV-infected C57BL/6 and BALB/c GM-BM, either in the absence or presence of LPS.

IL-15/IL-15Rα complex was produced at similar levels by C57BL/6 and BALB/c GM-BM, even after LPS-stimulation for 24 h (**Figure 7B**). ECTV infection significantly ($P > 0.01$) reduced the level of this complex in GM-BM from both mouse strains at 24 hpi. IL-18 was also secreted at comparable amounts by uninfected cells from both strains of mice, however, after stimulation with LPS for 24 h and ATP for additional 6 h, BALB/c GM-BM produced fourfold higher levels of IL-18 than C57BL/6 GM-BM (**Figure 7B**). ECTV-infection significantly ($P > 0.01$) inhibited TLR4 agonist + ATP-induced IL-18 production.

Because NO selectively enhances Th1 cell proliferation (Niedbala et al., 1999), we additionally checked the influence of ECTV infection on NO production by GM-BM from C57BL/6 and BALB/c mice (**Figure 7B**). Uninfected GM-BM from both mouse strains produced low levels of NO. After LPS treatment for 24 h, cells produced elevated levels of NO. In addition, after TLR4 agonist stimulation C57BL/6 GM-BM produced larger amounts of NO than BALB/c GM-BM. However, ECTV-infected cells from both mouse strains exhibited a significant ($P > 0.01$) decrease in NO production in response to LPS stimulation.

Additionally to Th1-polarizing cytokines, we evaluated the effect of ECTV infection on GM-BM capacity to produce cytokines that positively regulate Th2 polarization: IL-10, CCL2/MCP-1 and IL-1β, which has also Th1-polarizing properties (**Figures 7A,B**). IL-10 was undetectable in mock-, uvi-ECTV-, or ECTV-infected cultures of GM-BM from both mouse strains, indicating that ECTV is not able to stimulate IL-10 secretion by these cells (**Figure 7A**). After LPS treatment, uninfected BALB/c GM-BM produced twofold higher amounts of IL-10 than C57BL/6 cells. ECTV infection significantly ($P > 0.01$) decreased the production of IL-10 by LPS-treated GM-BM from both mouse strains. CCL2 was secreted at similar levels by C57BL/6 and BALB/c GM-BM, however, after LPS treatment cells from C57BL/6 mice produced slightly more CCL2 compared to BALB/c GM-BM (**Figure 7A**). ECTV infection significantly ($P > 0.01$) decreased the level of CCL2 secreted by GM-BM from both strains of mouse in response

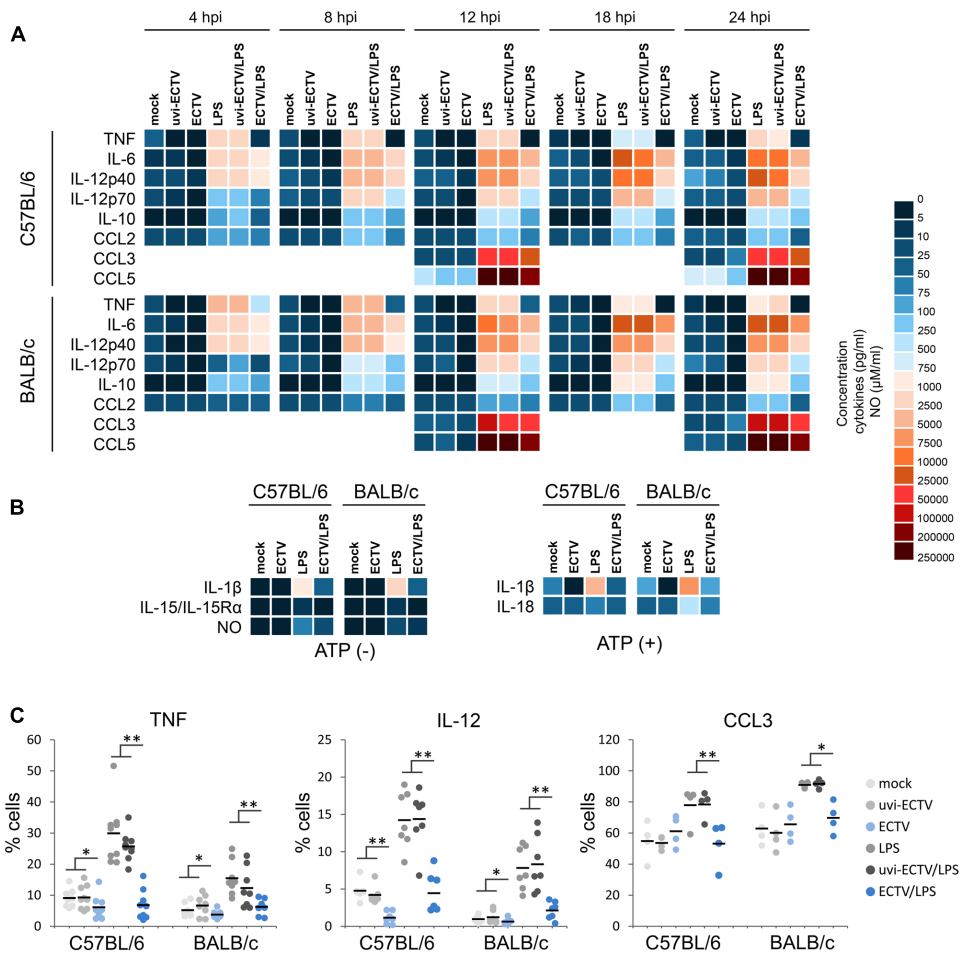


FIGURE 7 | The effect of ECTV infection on cytokine and chemokine synthesis and NO production by GM-BM is mouse strain-independent. Heat map represents mean concentration (pg/ml) values of cytokines and chemokines in culture supernatants of mock-, uvi-ECTV-, and ECTV-infected GM-BM, untreated or treated with LPS **(A)**, or additionally with ATP **(B)**. Mean NO concentration is presented as heat map ranging from 0 to 100 μM/ml **(B)**. **(C)** The percentage of cytokine-producing GM-BM at 24 hpi with ECTV. Each point represents an individual data, with the bar indicating the mean values (paired Student's *t*-test; $*P < 0.05$, $**P < 0.01$).

to LPS. Meanwhile, BALB/c GM-BM produced more IL-1β compared to C57BL/6 cells, especially after LPS treatment (**Figure 7B**). Further stimulation of cells with ATP additionally increased the level of secreted IL-1β. Similar to other cytokines, ECTV infection significantly ($P > 0.01$) reduced the level of IL-1β in supernatants of GM-BM from both mouse strains cultured under different conditions.

Intracellular staining of selected cytokines revealed that ECTV-infection also reduced the percentage of cells producing TNF-α and IL-12p40/p70 (**Figure 7C**). After LPS treatment the percentage of GM-BM from both mouse strains producing TNF-α, IL-12p40/p70, and CCL3 was significantly ($P > 0.05$) decreased. Uninfected GM-BM from C57BL/6 and BALB/c mice produced more TNF-α and IL-12p40/p70 or CCL3, respectively, in response to LPS treatment. Collectively, our data indicate that C57BL/6 and BALB/c GM-BM are able to produce higher amounts of Th1- and Th2-polarizing cytokines, respectively, especially in response to TLR4 agonist stimulation. ECTV infection causes a profound inhibition of the production of both Th1 and Th2-polarizing cytokines by GM-BM in a strain-independent manner, indicating that there is no differences in the reactivity of cDCs from resistant C57BL/6 and susceptible BALB/c mice to the virus.

The *in Vitro* Effect of ECTV-Infection on the Expression of MHC and Costimulatory Molecules on GM-BM Is Mouse Strain-Independent

We compared the influence of ECTV infection on the expression of MHC class I and II molecules, as well as costimulatory molecules, including CD40, CD80, and CD86, and CD83, on GM-BM generated from bone marrow progenitor cells of different mouse strains (**Figure 8**). GM-BM derived from C57BL/6 mice expressed lower levels of H-2Db molecules than BALB/c cells H-2Dd molecules. However, C57BL/6 cells expressed higher level of I-A/I-E, CD40 and CD80 molecules

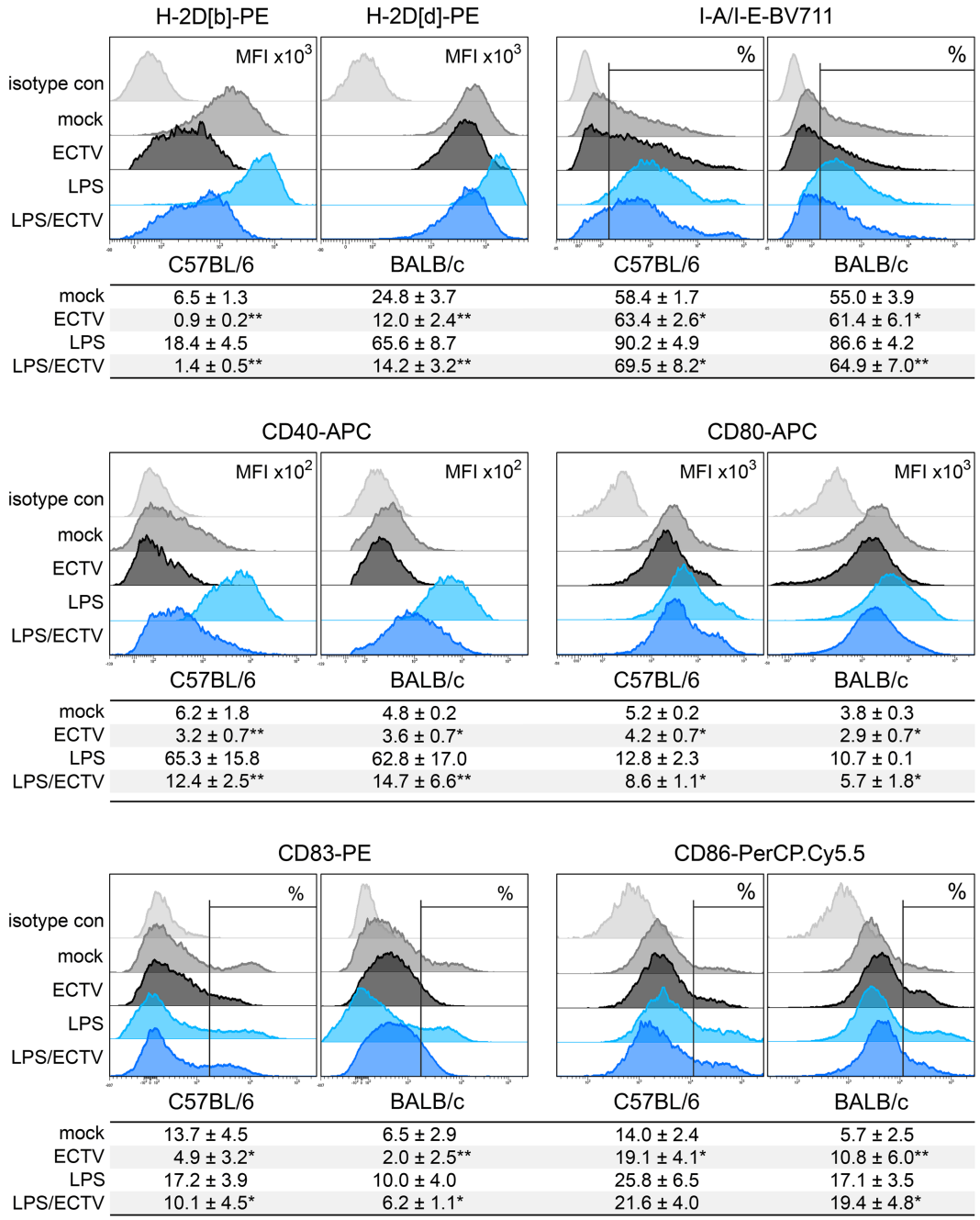


FIGURE 8 | ECTV infection inhibits maturation of GM-BM in a mouse strain-independent manner. Representative flow cytometry histograms showing the effect of ECTV infection and/or LPS treatment on maturation marker expression in GM-BM from C57BL/6 and BALB/c mice at 24 hpi. Below each histogram set is a table with mean values (±SD) for a given marker from at least three independent experiments. The expression is presented as mean fluorescent intensity (MFI) or the percentage of positive cells for a given marker (paired Student's *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$).

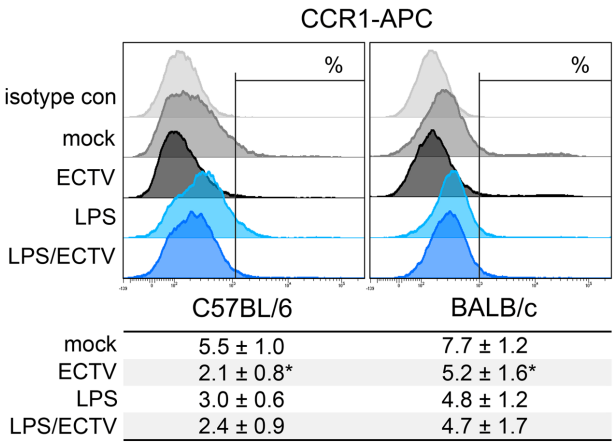
than those from BALB/c mice. Moreover, a higher percentage of GM-BM expressing CD83 and CD86 molecules was found in GM-BM derived from C57BL/6 than BALB/c mice (**Figure 8**). This suggests that GM-BM derived from C57BL/6 mice mature more efficiently than those from BALB/c mice. LPS treatment increased the expression of all tested molecules on GM-BM from both mouse strains in a maturation-dependent manner (**Figure 8**).

ECTV infection modulated the expression of all tested maturation molecules on GM-BM from both mouse strains. Infected cells from C57BL/6 and BALB/c mice showed significantly ($P > 0.01$) reduced mean fluorescent intensity

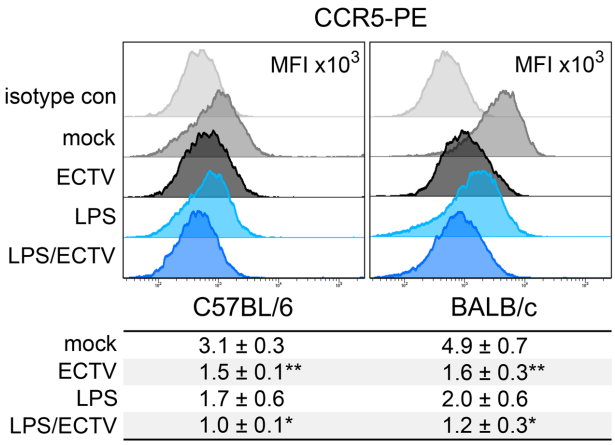
(MFI) of H-2Db and H-2Dd molecules, respectively, even after LPS treatment (**Figure 8**). The level of MFI for H-2D molecules was significantly more decreased in infected C57BL/6 compared to infected BALB/c cells, untreated (sevenfold vs. twofold; $P = 0.0015$) or treated (15-fold vs. 5-fold; $P = 0.0024$) with LPS. Meanwhile, in ECTV-infected cells we observed a significant ($P \leq 0.05$) increase in the percentage of I-A/I-E-positive cells, however, after LPS stimulation the percentage of such cells was reduced compared to mock-infected LPS-treated cells. Additionally, ECTV infection significantly ($P \leq 0.05$) reduced the expression of CD40 and CD80 costimulatory molecules on GM-BM from both mouse strains, even in the presence of LPS. Moreover, upon infection the percentage of $CD83^+$ cells significantly ($P \leq 0.05$) decreased in C57BL/6 and BALB/c GM-BM cultures. On the contrary, the percentage of $CD86^+$ cells was significantly ($P \leq 0.05$) increased in ECTV-exposed compared to mock-exposed cells from both mouse strains and in BALB/c cells treated with ECTV + LPS compared to cells treated only with LPS (**Figure 8**).

Chemokine Receptor Expression in Down-Regulated in ECTV-Infected GM-BM from C57BL/6 and BALB/c GM-BM

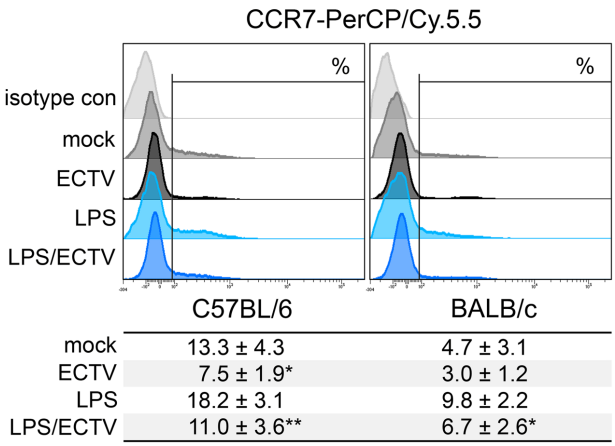
Our last question concerned the effect of ECTV infection on the expression of cell surface chemokine receptors that are differentially regulated upon maturation. CCR1 and CCR5 are reported to be reduced, and CCR7 is reported to be up-regulated on the surface of mature DCs (Le Nouën et al., 2011). Our results show that GM-BM from BALB/c mice expressed higher levels of chemokine receptors characteristic for immature DCs, compared to C57BL/6 cells (**Figure 9**). BALB/c cultures contained a higher percentage of $CCR1^+$ cells and expression of CCR5 on their surface was increased. On the contrary, C57BL/6 cultures had a higher percentage of $CCR7^+$ GM-BM. After infection with ECTV, cells from both mouse strains exhibited reduced percentage of $CCR1^+$ and $CCR7^+$ cells and decreased expression of CCR5, compared to mock-infected cells. After LPS treatment, the expression of chemokine receptors changed in a maturation dependent manner, i.e., the percentage of $CCR1^+$ and $CCR7^+$ cells decreased and increased, respectively, and MFI for CCR5 was low. Meanwhile, stimulation of ECTV-infected GM-BM with LPS resulted in the reduction of the percentage of $CCR7^+$ cells and MFI for CCR5 expression, compared to LPS-treated cells. Taken together, our results indicate that ECTV impairs expression of chemokine receptors on GM-BM and, therefore, it is not excluded that their potential to respond to chemokines regulating their migration is limited.



	C57BL/6	BALB/c
mock	5.5 ± 1.0	7.7 ± 1.2
ECTV	2.1 ± 0.8*	5.2 ± 1.6*
LPS	3.0 ± 0.6	4.8 ± 1.2
LPS/ECTV	2.4 ± 0.9	4.7 ± 1.7



	C57BL/6	BALB/c
mock	3.1 ± 0.3	4.9 ± 0.7
ECTV	1.5 ± 0.1**	1.6 ± 0.3**
LPS	1.7 ± 0.6	2.0 ± 0.6
LPS/ECTV	1.0 ± 0.1*	1.2 ± 0.3*



	C57BL/6	BALB/c
mock	13.3 ± 4.3	4.7 ± 3.1
ECTV	7.5 ± 1.9*	3.0 ± 1.2
LPS	18.2 ± 3.1	9.8 ± 2.2
LPS/ECTV	11.0 ± 3.6**	6.7 ± 2.6*

FIGURE 9 | ECTV infection similarly affects the expression of chemokine receptors on GM-BM from C57BL/6 and BALB/c mice. Representative flow cytometry histograms of chemokine receptor expression on GM-BM following the infection with ECTV and/or treatment with LPS at 24 hpi. Below each histogram set is a table with mean values (±SD) for a given marker from at least three independent experiments. The expression is presented as mean fluorescent intensity (MFI) or the percentage of positive cells for a given marker. Statistical comparisons were between mock- and ECTV-exposed DCs and between LPS- and ECTV + LPS-exposed DCs (paired Student's *t*-test; $^*P < 0.05$, $^{**}P < 0.01$).

DISCUSSION

The inter-strain differences in the reactivity of DCs to the antigen exposure/infection contributes to various types of adaptive immune responses and may partially determine the outcome of a disease (Liu et al., 2000, 2002). Therefore, in the

present study, we investigated the differences in the response of C57BL/6 and BALB/c GM-BM to ECTV infection by analysis of their maturation degree, chemokine receptor expression and production of different types of cytokines/chemokines engaged in T cell activation, polarization, and function. In addition, we analyzed the expression of key genes involved in maturation and activation of DCs. We found that *in vitro* ECTV modulation of GM-BM innate and adaptive immune properties occurred independently of the mouse strain susceptibility to infection. In cells from both strains of mice, ECTV infection contributed to the profound suppression of *polarizing signals to prime* Th1 immune response. Furthermore, contrary to expectations, a more pronounced viral inhibitory effect was observed in resistant C57BL/6 GM-BM, despite their higher potential to stimulate Th1 response under physiological conditions. These results indicate strong adaptation capacity of ECTV to the natural host DCs and its direct inhibitory effect on DC properties is irrespective of whether they are derived from resistant or susceptible mouse strain.

Several *in vivo* studies have reported that Th1 and Th2 immune responses prevail in C57BL/6 and BALB/c mice, respectively, under physiological conditions (Trunova et al., 2011), stress (Palumbo et al., 2010), or infection (Belkaid et al., 2002; Sacks and Noben-Trauth, 2002; Chaudhri et al., 2004; Watanabe et al., 2004). The immune response of different types of cells from these mouse strains, including DCs and macrophages, may influence the development of Th1 and Th2 adaptive immunity (Liu et al., 2002; Watanabe et al., 2004). It has been shown that DCs from C57BL/6 mice produce higher levels of IL-12 and IL-15 than DCs from BALB/c mice during early stages of *Listeria monocytogenes* infections (Liu et al., 2000). Moreover, DCs isolated from spleens of naïve C57BL/6 mice preferentially expressed a higher level of TLR9 mRNA, whereas those from BALB/c mice expressed a higher level of mRNAs for TLR2, -4, -5, and -6. In response to microbial ligands for TLR2 (lipoprotein), TLR4 (LPS), TLR2/6 (zymosan), and TLR9 (CpG), DCs from C57BL/6 and BALB/c mice produced larger amounts of IL-12p40 and CCL2, respectively. Additionally, DCs from C57BL/6 mice exhibited a more mature phenotype than BALB/c DCs due to higher expression of maturation markers, such as CD40 and CD86, and signal transducer and activator of transcription 4 (Stat4)—a key intermediate in IL-12 signaling pathway and Th1 cell development (Liu et al., 2002).

Our *in vitro* studies revealed that GM-BM derived from C57BL/6 mice displayed higher expression of *Ccl5*, *Cd40*, *Il-12b*, *Irf7*, *Thbs1*, and *Tlr9* than BALB/c cells. Moreover, upon LPS stimulation, *Cd80*, *Cd86*, *Ccl5*, and *Tlr9* were more up-regulated in C57BL/6 than BALB/c GM-BM. Additionally, C57BL/6 cells produced higher levels of IL-12p40 and IL-12p70, whereas BALB/c cells secreted more TNF-α, IL-6, and IL-10, especially after LPS treatment. Higher expression of maturation markers, such as MHC II, CD40, and CD80, and large percentage of cells expressing CD83, CD86, and CCR7 was observed in C57BL/6 cultures. Overall, our data indicate that GM-BM from resistant C57BL/6 mice are more matured than cells from susceptible BALB/c cells and possess a higher potential to stimulate Th1 response. Inter-strain differences in maturation state were also observed between GM-BM from BALB/c, C57BL/6, and atopic prone NC/Nga mice (Koike et al., 2008). The latter had greater expression of MHC class II, CD80, CD86, and CD11c compared to cells from BALB/c and C57BL/6 mice, as well as preferentially stimulated Th2 cytokine immune response. Therefore, it is suggested that the genetic background may enhance the differentiation and function of DCs and may be partially related to the development or aggravation of allergic/atopic diseases (Koike et al., 2008). On the contrary, resting BMDCs generated with GM-CSF and IL-4 from bone marrow progenitor cells of C57BL/6 and BALB/c mouse strains displayed no differences in the expression of CD40, CD80, CD86, MHC II, and TLR2 (Jiang et al., 2010).

It has been proposed that differences in innate and adaptive immune functions of DCs underline additional mechanisms responsible for resistance or susceptibility of C57BL/6 or BALB/c mice to *L. monocytogenes* (Liu et al., 2000, 2002). Moreover, DC capacity to differentiate naive T cells into functional Th1 or Th2 effector cells during *Leishmania major* infection is cell-intrinsic and Th2 polarization is not restricted to H2-d mice, since DCs from BALB/c and B10.D2 DCs (both H2d) exhibited differences in maturation and ability to induce Th differentiation (Filippi et al., 2003). In our *in vitro* study, the influence of ECTV infection on C57BL/6 and BALB/c GM-BM functions was comparable. Cells from both strains of mice exhibited a profound immunosuppression due to the productive virus replication (Szulc-Dąbrowska et al., 2017). Interestingly, GM-BM from resistant C57BL/6 mice had a higher percentage of $ECTV^+$ cells compared to BALB/c cells at 24 hpi, suggesting that the viral spread is more efficient in these cells at later stages of infection. Moreover, upon ECTV infection 11 genes (*Rac1*, *Ccl2*, *Cxcl12*, *Mif*, *Ccr1*, *Cxcr4*, *Lyn*, *Fcgr1*, *Cebpa*, *Irf7*, and *Stat3*) were significantly more repressed in C57BL/6 cells compared to BALB/c cells, whereas the latter only had two genes (*Ccl3* and *Cd74*) more down-regulated compared to C57BL/6 cells. Additionally, infected GM-BM from C57BL/6 mice displayed a more expressive reduction in MHC I expression than BALB/c GM-BM. It is not excluded that slightly increased virus inhibitory effect on C57BL/6 GM-BM function is associated with a higher rate of ECTV dissemination within these cells. The ability of ECTV to replicate productively in GM-BM is a manifestation of its high adaptation capacity to the natural host immune cells, since other orthopoxviruses, such as CPXV and VACV abortively infect human DCs (Engelmayer et al., 1999; Jenne et al., 2000; Hansen et al., 2011).

Our studies revealed that ECTV infection in GM-BM derived from C57BL/6 and BALB/c mice down-regulated many genes involved in antigen uptake and processing, chemokine and cytokine synthesis, receptor expression, and signal transduction. The inhibitory effect on gene expression was also observed after TLR4 agonist treatment. In general, the analyzed genes were regulated in the same way in cells from both strains of mice and we did not observe any strain-specific response in GM-BM upon *in vitro* ECTV-infection. On the contrary, peritoneal macrophages from C57BL/6 and BALB/c mice infected with ECTV exhibited differential regulation of 14 innate antiviral genes, which were up-regulated in C57BL/6 and down-regulated

in BALB/c cells, suggesting that these variations in gene expression may partially contribute to resistance or susceptibility to severe mousepox (Dolega et al., 2017). Meantime, several inter-strain variations in GM-BM functionality have been observed during paramyxovirus simian virus 5 (SV5) infection (Pejawar et al., 2005). Firstly, GM-BM from C57BL/6 mice were much more permissive to SV5 infection than cells derived from BALB/c mice. Secondly, despite the production of a similar panel of cytokines, cells differed in the maturation state: C57BL/6 cells up-regulated the expression of CD40, CD80, and CD86, whereas BALB/c cells displayed increase only in CD40 and CD86 expression. Thirdly, SV5-matured C57BL/6 GM-BM were more potent to activate naïve $CD8^+$ T cells than SV5-matured BALB/c cells (Pejawar et al., 2005). Similar inter-strain differences were observed in BMDCs during infection with *Chlamydia muridarum* (Jiang et al., 2010). After *in vitro* infection, BMDCs from C57BL/6 mice underwent higher functional maturation than cells from BALB/c mice. This was reflected by higher expression of MHC class II and costimulatory molecules (CD40, CD80, and CD86) and greater production of IL-12. On the contrary, BALB/c BMDCs secreted more IL-23, IL-6, IL-10, and TNF-α than C57BL/6 cells (Jiang et al., 2010). Overall, the data described above demonstrate that genetically defined differences in functionality between DCs from C57BL/6 and BALB/c mice may be phenotypically expressed during exposure to the microbial agent/infection, however, this reactivity also may depend on the nature of the agent and the host–agent immunobiology.

ECTV infection of GM-BM derived from C57BL/6 and BALB/c mice led to a profound repression of a set of chemokine and cytokine genes involved in Th1 [*Ccl3*, *Ccl4*, *Ccl5* (Lebre et al., 2005), *Cxcl2* (Seow et al., 2008), *Cxcl10* (Lebre et al., 2005), *Ifng*, *Il12a*, *Il12b*, and *Il16* (Lynch et al., 2003)] and Th2 [*Ccl2* (Liu et al., 2002), *Ccl8* (Lech and Anders, 2013), *Ccl11* (Dixon et al., 2006), *Ccl17* (Belperio et al., 2004), *Cxcl12* (Piao et al., 2012), *Mif* (Das et al., 2011), and *Tgfb1* (Maeda and Shiraishi, 1996)] immune response regulation. But, ECTV infection up-regulated mRNA transcript for IL-10 in GM-BM of both strains of mice. Despite the increase in gene expression, the level of IL-10 remained undetectable in the culture supernatants during 24 h incubation. We cannot exclude that GM-BM were cultured at too low cell density to detect IL-10 in our experimental model. On the other hand, it is possible that ECTV does not stimulate IL-10 production by GM-BM, similar to CPXV, which has been shown not to induce the secretion of IL-10 by different types of human DCs: monocyte-derived DCs, myeloid DCs, and plasmacytoid DCs (Hansen et al., 2011).

IL-10, first described as a Th2-polarizing cytokine, is an anti-inflammatory cytokine that suppresses pro-inflammatory response leading to increased pathogen dissemination and/or reduced pathology. IL-10 secreted by DCs may act in an autocrine manner to inhibit production of chemokines (such as CCL2, CCL5, CCL12, CXCL8, and CXCL10) and pro-inflammatory cytokines (such as IL-1α/β, IL-6, IL-12, IL-18, and TNF-α). Moreover, IL-10 may directly inhibit proliferation and production of IL-2, IFN-γ, IL-4, IL-5, and TNF-α by $CD4^+$ T cells, thus it regulates both Th1 and Th2 immune responses (Couper et al., 2008). Additionally, IL-10 induces long-lasting T cell anergy and promotes the differentiation of naïve T cells into Tr1 cells in humans and mice (Gregori et al., 2010). The importance of cellular IL-10 in immune system regulation is supported by the fact that several viruses, including seven members of the *Poxviridae* family, encode orthologs of cellular IL-10, called viral IL-10s (vIL-10s), which have been acquired by viruses from their host during evolution (Ouyang et al., 2014).

It has been shown that CPXV induces *in vitro* secretion of IL-10 by BMDCs and RAW 264.7 macrophages at 24 hpi. Moreover, CPXV is able to induce higher IL-10 production *in vivo* than VACV (Spesock et al., 2011). Experiments with IL-10-deficient mice have indicated that after intranasal CPXV infection these mice exhibited similar weight loss and viral burdens as wild-type mice. However, IL-10-deficient mice were more susceptible to CPXV reinfection, because increased viral loads were observed in their lungs, what corresponded with lower antibody and $CD8^+$ T cell responses compared to wild-type mice. The role of IL-10 during CPXV infection is probably beneficial for the virus, and IL-10 suppresses immunopathology in the lungs because IL-10-deficient mice after re-challenge with CPXV displayed greater bronchopneumonia than wild-type mice (Spesock et al., 2011). Meanwhile, using recombinant VACV expressing mouse IL-10 (mIL-10) it has been shown that in immunocompetent mice mIL-10 expressed from the VACV genome affected natural killer (NK) and virus-specific CTL activity, whereas in severe combined immunodeficient (SCID) mice VACV-mIL-10 infection resulted in increased NK cell activity and higher degree of virus clearance compared to infection with control VACV (Kurilla et al., 1993). In agreement with these studies, van Den Broek et al. (2000) have demonstrated that IL-10-deficient mice showed dramatically reduced viral titer in ovaries after intraperitoneal infection, suggesting a rapid viral clearance. Together, these data indicate that IL-10 is a potent cytokine in suppressing the immune response against VACV and may be a dominant factor for susceptibility to acute VACV infection (van Den Broek et al., 2000). On the hand, more recently it has been found, using multiphoton intravital microscopy imaging, that IL-10 produced locally in the skin of VACV-infected mice contributed to viral clearance, probably by shaping the innate immune response within the inflamed tissue and/or by reducing virus-induced inflammation (Cush et al., 2016). The up-regulation of IL-10 mRNA level observed in ECTV-infected GM-BM may therefore suggest that these cells possess a stronger ability to damp of the immune response and/or cross-regulate Th1 and Th2 immune responses. It is highly possible that IL-10 may be an important factor in the generation of non-protective Th2 immune response in susceptible BALB/c mice *in vivo*.

Our studies revealed that ECTV infection impairs inflammatory response and maturation of GM-BM in a strain-independent manner. At 24 hpi GM-BM from C57BL/6 and BALB/c mice displayed an altered production of several cytokines and chemokines with the exception of CCL3, which was produced at a higher concentration. Interestingly, mRNA transcript for CCL3 was down-regulated at this time point in ECTV-infected cells. It is known that mRNA expression does not always correlate with protein levels in mammalian

cells (Vogel and Marcotte, 2012). Moreover, ECTV infection of GM-BM from both mouse strains down-regulated expression of MHC class I, and CD40 and CD80 co-stimulatory molecules, but increased the percentage of MHC II^+ and $CD86^+$ cells. Possibly, bystander non-infected GM-BM underwent partial maturation at the same time being the main source of CCL3. It has been demonstrated that highly attenuated modified vaccinia virus Ankara (MVA) induces phenotypic and functional maturation of bystander DCs resulting in production of a large array of cytokines and chemokines involved in T cell activation and recruitment, and regulation of inflammatory response, including CCL3 (Pascutti et al., 2011). However, ECTV encodes vCCI decoy receptor, EVM1, which is an abundantly secreted glycoprotein during infection and can bind the CC chemokines to form highly stable complexes with CCL3 and CCL5 (Arnold and Fremont, 2006). Our results also showed that the inhibition of inflammatory response and maturation of C57BL/6 and BALB/c GM-BM caused by ECTV was most pronounced after LPS treatment. This observation is in agreement with previous *in vitro* studies showing that other orthopoxviruses, such as VACV and CPXV severely affected maturation and activation of cDCs (Engelmayer et al., 1999; Jenne et al., 2000; Hansen et al., 2011).

As a master of immune inhibitory strategies, ECTV also altered the expression of chemokine receptors, such as CCR1, CCR5, and CCR7 on the C57BL/6 and BALB/c GM-BM surface, what may lead to their impaired migration (Jang et al., 2006). Humrich et al. (2007) have demonstrated that VACV infection targets chemokine-induced migration of DCs at multiple functional levels. VACV-infected mature DCs showed inability to migrate toward the lymphoid chemokines CCL19 and CXCL12 without apparent alterations in expression of surface chemokine receptors CXCR4 and CCR7. In fact, in VACV-infected immature or uninfected bystander DCs there is decreased or increased, respectively, expression of the inflammatory chemokine receptors CCR1 and CXCR1, which abrogates or intensifies their migration toward CCL3 and CCL5. Moreover, VACV-infected and uninfected bystander DCs are not able to up-regulate CCR7 expression after LPS treatment suggesting their disability to undergo chemokine receptor switch (Humrich et al., 2007). Additionally, poxviruses encode membrane cytokine and chemokine receptors to evade host immune responses (Felix and Savvides, 2017).

Taken together, our results indicate that *in vitro* ECTV infection of GM-BM, including cDCs, leads to their functional impairment independently of the genetic background of mice from which they were generated. ECTV-employed host-specific strategies to evade host antiviral immune response allow the virus to control GM-BM independently of the host resistance or susceptibility to severe mousepox. Moreover, our study confirms that ECTV is a master of immune inhibitory strategies showing a wide array of mechanisms for disrupting the innate and acquired immune functions of GM-BM. Better understanding of the virus interactions with cDCs, the most potent APCs, may help to elucidate additional mechanisms responsible for resistance or susceptibility to mousepox and that can lead to rational design means and ways of containing virus infection.

AUTHOR CONTRIBUTIONS

LS-D conceived and designed the study. LS-D, JC, ZN, and FT conducted real-time PCR experiments. LS-D, AW, and MG performed flow cytometry analysis. LS-D and JS performed ELISA. LS-D, JS, FT, and MG analyzed and interpreted the data. LS-D prepared figures and wrote the draft of the manuscript. All authors reviewed and approved the manuscript.

ACKNOWLEDGMENTS

This work was supported by grant No. UMO-2012/05/D/NZ6/02916 to LS-D from the National Science Center in Poland. Results were partially presented at the 17th International Congress of Immunology—ICI 2016, 21–26 August 2016, Melbourne, VIC, Australia, in which LS-D's participation was supported by KNOW (Leading National Research Centre) Scientific Consortium "Healthy Animal – Safe Food," decision of Ministry of Science and Higher Education No. 05-1/KNOW2/2015.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2017.02539/full#supplementary-material

FIGURE S1 | Characteristics of C57BL/6 and BALB/c GM-BM before and after MACS separation of $CD11c^+$ cells. **(A)** Representative dot plot demonstrating gating strategy of isotype controls and CD11c and I-A/I-E staining. **(B)** Representative histograms demonstrating MFI of CD11c, CD11b, and CD205 expression on C57BL/6 and BALB/c GM-BM before and after MACS separation.

TABLE S1 | Gene list description of Mouse Dendritic and Antigen Presenting Cell RT2 Profiler PCR Array (Qiagen).

REFERENCES

Arnold, P. L., and Fremont, D. H. (2006). Structural determinants of chemokine binding by an ectromelia virus-encoded decoy receptor. *J. Virol.* 80, 7439–7449. doi: 10.1128/JVI.00576-06

Belkaid, Y., Piccirillo, C. A., Mendez, S., Shevach, E. M., and Sacks, D. L. (2002). CD4+ CD25+ regulatory T cells control *Leishmania* major persistence and immunity. *Nature* 420, 633–637.

Belperio, J. A., Dy, M., Murray, L., Burdick, M. D., Xue, Y., Strieter, R. M., et al. (2004). The role of the Th2 CC chemokine ligand CCL17 in pulmonary fibrosis. *J. Immunol* 173, 4692–4698. doi: 10.4049/jimmunol.173.7.4692

Brownstein, D. G., and Gras, L. (1995). Chromosome mapping of Rmp-4, a gonad-dependent gene encoding host resistance to mousepox. *J. Virol.* 69, 6958–6964.

Chaudhri, G., Panchanathan, V., Buller, R. M. L., van den Eertwegh, A. J. M., Claassen, E., Zhou, J., et al. (2004). Polarized type 1 cytokine response and

cell-mediated immunity determine genetic resistance to mousepox. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9057–9062. doi: 10.1073/pnas.0402949101

Couper, K. N., Blount, D. G., and Riley, E. M. (2008). IL-10: the master regulator of immunity to infection. *J. Immunol.* 180, 5771–5777. doi: 10.4049/jimmunol.180.9.5771

Cush, S. S., Reynoso, G. V., Kamenyeva, O., Bennink, J. R., Yewdell, J. W., and Hickman, H. D. (2016). Locally produced IL-10 limits cutaneous *Vaccinia virus* spread. *PLOS Pathog.* 12:e1005493. doi: 10.1371/journal.ppat.1005493

Das, R., Moss, J. E., Robinson, E., Roberts, S., Levy, R., Mizue, Y., et al. (2011). Role of macrophage migration inhibitory factor in the Th2 immune response to epicutaneous sensitization. *J. Clin. Immunol.* 31, 666–680. doi: 10.1007/s10875-011-9541-7

Dixon, H., Blanchard, C., Deschoolmeester, M. L., Yuill, N. C., Christie, J. W., Rothenberg, M. E., et al. (2006). The role of Th2 cytokines, chemokines and parasite products in eosinophil recruitment to the gastrointestinal mucosa during helminth infection. *Eur. J. Immunol.* 36, 1753–1763. doi: 10.1002/eji.200535492

Dolega, P., Szulc-Dąbrowska, L., Bossowska, M., Mielcarska, M., Nowak, Z., and Toka, F. N. (2017). Innate immune gene transcript level associated with the infection of macrophages with ectromelia virus in two different mouse strains. *Viral. Immunol.* 5, 315–329. doi: 10.1089/vim.2016.0173

Engelmayer, J., Larsson, M., Subklewe, M., Chahroudi, A., Schmaljohn, A., William, C., et al. (1999). *Vaccinia virus* inhibits the maturation of human dendritic cells: a novel mechanism of immune evasion. *J. Immunol.* 163, 6762–6768.

Felix, J., and Savvides, S. N. (2017). Mechanisms of immunomodulation by mammalian and viral decoy receptors: insights from structures. *Nat. Rev. Immunol.* 17, 112–129. doi: 10.1038/nri.2016.134

Filippi, C., Hugues, S., Cazareth, J., Julia, V., Glaichenhaus, N., and Ugolini, S. (2003). CD4+ T cell polarization in mice is modulated by strain-specific major histocompatibility complex–independent differences within dendritic cells. *J. Exp. Med.* 198, 201–209. doi: 10.1084/jem.20021893

Gregori, S., Tomasoni, D., Pacciani, V., Scirpoli, M., and Battaglia, M. (2010). Differentiation of type 1 T regulatory cells (Tr1) by tolerogenic DC-10 requires the IL-10-dependent ILT4/HLA-G pathway. *Blood* 116, 935–944. doi: 10.1182/blood-2009-07-234872

Hansen, S. J., Rushton, J., Dekonenko, A., Chand, H. S., Olson, G. K., Hutt, J. A., et al. (2011). Cowpox virus inhibits human dendritic cell immune function by nonlethal, nonproductive infection. *Virology* 412, 411–425. doi: 10.1016/j.virol.2011.01.024

Humrich, J. Y., Thumann, P., Greiner, S., Humrich, J. H., Averbeck, M., Schwank, C., et al. (2007). *Vaccinia virus* impairs directional migration and chemokine receptor switch of human dendritic cells. *Eur. J. Immunol.* 37, 954–965. doi: 10.1002/eji.200636230

Jang, M. H., Sougawa, N., Tanaka, T., Hirata, T., Hiroi, T., Tohya, K., et al. (2006). CCR7 is critically important for migration of dendritic cells in intestinal lamina propria to mesenteric lymph nodes. *J. Immunol.* 176, 803–810. doi: 10.4049/jimmunol.176.2.803

Jenne, L., Hauser, C., Arrighi, J. F., Saurat, J. H., and Hugin, A. W. (2000). Poxvirus as a vector to transduce human dendritic cells for immunotherapy: abortive infection but reduced APC function. *Gene Ther.* 7, 1575–1583.

Jiang, X., Shen, C., Yu, H., Karunakaran, K. P., and Brunham, R. C. (2010). Differences in innate immune responses correlate with differences in murine susceptibility to *Chlamydia muridarum* pulmonary infection. *Immunology* 129, 556–566. doi: 10.1111/j.1365-2567.2009.03157.x

Kaiko, E. G., Horvat, J. C., Beagley, K. W., and Hansbro, P. M. (2008). Immunological decision-making: how does the immune system decide to mount a helper T-cell response? *Immunology* 123, 326–338. doi: 10.1111/j.1365-2567.2007.02719.x

Koike, E., Takano, H., Inoue, K., and Yanagisawa, R. (2008). Accelerated differentiation of bone marrow-derived dendritic cells. *Int. Immunopharmacol.* 8, 1737–1743. doi: 10.1016/j.intimp.2008.08.006

Kurilla, M. G., Swaminathan, S., Welsh, R. M., Kieff, E., and Brutkiewicz, R. R. (1993). Effects of virally expressed interleukin-10 on *Vaccinia virus* infection in mice. *J. Virol.* 67, 7623–7628.

Le Nouën, C., Hillyer, P., Winter, C. C., McCarty, T., Rabin, R. L., Collins, P. L., et al. (2011). Low CCR7-mediated migration of human monocyte derived dendritic cells in response to human respiratory syncytial virus and human metapneumovirus. *PLOS Pathog.* 7:e1002105. doi: 10.1371/journal.ppat.1002105

Lebre, M. C., Burwell, T., Vieira, P. L., Lora, J., Coyle, A. J., Kapsenberg, M. L., et al. (2005). Differential expression of inflammatory chemokines by Th1- and Th2-cell promoting dendritic cells: a role for different mature dendritic cell populations in attracting appropriate effector cells to peripheral sites of inflammation. *Immunol. Cell Biol.* 83, 525–535. doi: 10.1111/j.1440-1711.2005.01365.x

Lech, M., and Anders, H.-J. (2013). Macrophages and fibrosis: how resident and infiltrating mononuclear phagocytes orchestrate all phases of tissue injury and repair. *Biochim. Biophys. Acta* 1832, 989–997. doi: 10.1016/j.bbadis.2012.12.001

Liu, T., Matsuguchi, T., Tsuboi, N., Yajima, T., and Yoshikai, Y. (2002). Differences in expression of Toll-like receptors and their reactivities in dendritic cells in BALB/c and C57BL/6 mice. *Infect. Immun.* 70, 6638–6645. doi: 10.1128/IAI.70.12.6638-6645.2002

Liu, T., Nishimura, H., Matsuguchi, T., and Yoshikai, Y. (2000). Difference in interleukin-12 and -15 production by dendritic cells at the early stage of *Listeria monocytogenes* infection between BALB/c and C57BL/6 mice. *Cell. Immunol.* 202, 31–40. doi: 10.1128/IAI.70.12.6638-6645.2002

Lutz, M. B. (2016). Induction of CD4+ regulatory and polarized effector/helper T cells by dendritic cells. *Immune Netw.* 16, 13–25. doi: 10.4110/in.2016.16.1.13

Lutz, M. B., Kukutsch, N., Ogilvie, A. L., Rössner, S., Koch, F., Romani, N., et al. (1999). An advanced culture method for generating large quantities of highly pure dendritic cells from mouse bone marrow. *J. Immunol. Methods* 223, 77–92. doi: 10.1016/S0022-1759(98)00204-X

Lynch, E. A., Heijens, C. A., Horst, N. F., Center, D. M., and Cruikshank, W. W. (2003). Cutting edge: IL-16/CD4 preferentially induces Th1 cell migration: requirement of CCR5. *J. Immunol.* 171, 4965–4968. doi: 10.4049/jimmunol.171.10.4965

Maeda, H., and Shiraishi, A. (1996). TGF-beta contributes to the shift toward Th2-type responses through direct and IL-10-mediated pathways in tumor-bearing mice. *J. Immunol.* 156, 73–78.

McCollum, A. M., and Damon, I. K. (2014). Human monkeypox. *Clin. Infect. Dis.* 58, 260–267. doi: 10.1093/cid/cit703

Mehta, V. B., Hart, J., and Wewers, D. (2000). ATP stimulated release of IL-1β and IL-18 requires priming by LPS and is independent of caspase-1 cleavage. *J. Biol. Chem.* 276, 3820–3826. doi: 10.1074/jbc.M006814200

Niedbala, W., Wei, X. Q., Piedrafita, D., Xu, D., and Liew, F. Y. (1999). Effects of nitric oxide on the induction and differentiation of Th1 cells. *Eur. J. Immunol.* 29, 2498–2505.

Ouyang, P., Rakus, K., van Beurden, S. J., Westphal, A. H., Davison, A. J., Gatherer, D., et al. (2014). IL-10 encoded by viruses: a remarkable example of independent acquisition of a cellular gene by viruses and its subsequent evolution in the viral genome. *J. Gen. Virol.* 95, 245–262. doi: 10.1099/vir.0.058966-0

Palumbo, M. L., Canzobre, M. C., Pascuan, C. G., Ríos, H., Wald, M., and Genaro, A. M. (2010). Stress induced cognitive deficit is differentially modulated in BALB/c and C57Bl/6 mice: correlation with Th1/Th2 balance after stress exposure. *J. Neuroimmunol.* 218, 12–20. doi: 10.1016/j.jneuroim.2009.11.005

Parker, S., Siddiqui, A. M., Painter, G., Schriewer, J., and Buller, R. M. (2010). Ectromelia virus infections of mice as a model to support the licensure of anti-orthopoxvirus therapeutics. *Viruses* 2, 1918–1932. doi: 10.3390/v2091918

Pasare, C., and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+ CD25+ T cell-mediated suppression by dendritic cells. *Science* 299, 1033–1036. doi: 10.1126/science.1078231

Pascutti, M. F., Rodriguez, A. M., Falivene, J., Giavedoni, L., Drexler, I., and Gherardi, M. M. (2011). Interplay between modified *Vaccinia virus* Ankara and dendritic cells: phenotypic and functional maturation of bystander dendritic cells. *J. Virol.* 85, 5532–5545. doi: 10.1128/JVI.02267-10

Pejawar, S. S., Parks, G. D., and Alexander-Miller, M. A. (2005). Abortive versus productive viral infection of dendritic cells with a paramyxovirus results in differential upregulation of select costimulatory molecules. *J. Virol.* 79, 7544–7557. doi: 10.1128/JVI.79.12.7544-7557.2005

Piao, H.-L., Tao, Y., Zhu, R., Wang, S.-C., Tang, C. L., Fu, Q., et al. (2012). The CXCL12/CXCR4 axis is involved in the maintenance of Th2 bias at the maternal/fetal interface in early human pregnancy. *Cell. Mol. Immunol.* 9, 423–430. doi: 10.1038/cmi.2012.23

Riedel, S. (2005). Smallpox and biological warfare: a disease revisited. *Proc. (Bayl. Univ. Med. Cent.)* 18, 13–20.

Russe, O. Q., Möser, C. V., Kynast, K. L., King, T. S., Olbrich, K., Grösch, S., et al. (2014). LPS inhibits caspase 3-dependent apoptosis in RAW264.7 macrophages induced by the AMPK activator AICAR. *Biochem. Biophys. Res. Commun.* 447, 520–525. doi: 10.1016/j.bbrc.2014.04.008

Sacks, D., and Noben-Trauth, N. (2002). The immunology of susceptibility and resistance to *Leishmania* major in mice. *Nat. Rev. Immunol.* 2, 845–858. doi: 10.1038/nri933

Sei, J. J., Haskett, S., Kaminsky, L. W., Lin, E., Truckenmiller, M. E., Bellone, C. J., et al. (2015). Peptide-MHC-I from endogenous antigen outnumber those from exogenous antigen, irrespective of APC phenotype or activation. *PLOS Pathog.* 11:e1004941. doi: 10.1371/journal.ppat.1004941

Seow, S. W., Rahmat, J. N., Bay, B. H., Lee, Y. K., and Mahendran, R. (2008). Expression of chemokine/cytokine genes and immune cell recruitment following the instillation of *Mycobacterium bovis*, *Bacillus* Calmette–Guérin or *Lactobacillus rhamnosus* strain GG in the healthy murine bladder. *Immunology* 124, 419–427. doi: 10.1111/j.1365-2567.2007.02792.x

Spesock, A. H., Barefoot, B. E., Ray, C. A., Kenan, D. J., Gunn, M. D., Ramsburg, E. A., et al. (2011). Cowpox virus induces interleukin-10 both *in vitro* and in vivo. *Virology* 417, 87–97. doi: 10.1016/j.virol.2011.05.010

Stanford, M. M., McFadden, G., Karupiah, G., and Chaudhri, G. (2007). Immunopathogenesis of poxvirus infections: forecasting the impending storm. *Immunol. Cell Biol.* 85, 93–102. doi: 10.1038/sj.icb.7100033

Szulc-Dabrowska, L., Gregorczyk, K. P., Struzik, J., Boratynska-Jasinska, A., Szczepanowska, J., Wyzewski, Z., et al. (2016). Remodeling of the fibroblast cytoskeletal architecture during the replication cycle of *Ectromelia virus*: a morphological *in vitro* study in a murine cell line. *Cytoskeleton* 7, 396–417. doi: 10.1002/cm.21308

Szulc-Dąbrowska, L., Struzik, J., Ostrowska, A., Guzera, M., Toka, F. N., Bossowska-Nowicka, M., et al. (2017). Functional paralysis of GM-CSF–derived bone marrow cells productively infected with *Ectromelia virus*. *PLOS ONE* 12:e0179166. doi: 10.1371/journal.pone.0179166

Trunova, G. V., Makarova, O. V., Diatroptov, M. E., Bogdanova, I. M., Mikchailova, L. P., and Abdulaeva, S. O. (2011). Morphofunctional characteristic of the immune system in BALB/c and C57Bl/6 mice. *Bull. Exp. Biol. Med.* 151, 99–102.

van Den Broek, M., Bachmann, M. F., Köhler, G., Barner, M., Escher, R., Zinkernagel, R., et al. (2000). IL-4 and IL-10 antagonize IL-12-mediated protection against acute *Vaccinia virus* infection with a limited role of IFN-gamma and nitric oxide synthetase 2. *J. Immunol.* 164, 371–378. doi: 10.4049/jimmunol.164.1.371

Vogel, C., and Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13, 227–232. doi: 10.1038/nrg3185

Watanabe, K., Numata, K., Ito, T., Takagi, K., and Matsukawa, A. (2004). Innate immune response in Th1- and Th2-dominant mouse strains. *Shock* 22, 460–466. doi: 10.1097/01.shk.0000142249.08135.e9

Zhang, W., Chen, Z., Li, F., Kamencic, H., Juurlink, B., Gordon, J. R., et al. (2003). Tumour necrosis factor-α (TNF-α) transgene-expressing dendritic cells (DCs) undergo augmented cellular maturation and induce more robust T-cell activation and anti-tumour immunity than DCs generated in recombinant TNF-α. *Immunology* 108, 177–188. doi: 10.1046/j.1365-2567.2003.01489.x

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

The reviewer RALR and handling Editor declared their shared affiliation.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 March 2018
doi: 10.3389/fmicb.2018.00402

Serological Evidence of *Orthopoxvirus* Circulation Among Equids, Southeast Brazil

Iara A. Borges[1], Mary G. Reynolds[2], Andrea M. McCollum[2], Poliana O. Figueiredo[1], Lara L. D. Ambrosio[1], Flavia N. Vieira[1], Galileu B. Costa[1], Ana C. D. Matos[1], Valeria M. de Andrade Almeida[1], Paulo C. P. Ferreira[1], Zélia I. P. Lobato[1], Jenner K. P. dos Reis[1], Erna G. Kroon[1] and Giliane S. Trindade[1]*

[1] Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [2] Centers for Disease Control and Prevention (CDC), Atlanta, GA, United States

Since 1999 *Vaccinia virus* (VACV) outbreaks involving bovines and humans have been reported in Brazil; this zoonosis is known as Bovine Vaccinia (BV) and is mainly an occupational disease of milkers. It was only in 2008 (and then again in 2011 and 2014) however, that VACV was found causing natural infections in Brazilian equids. These reports involved only equids, no infected humans or bovines were identified, and the sources of infections remain unknown up to date. The peculiarities of Equine Vaccinia outbreaks (e.g., absence of human infection), the frequently shared environments, and fomites by equids and bovines in Brazilian farms and the remaining gaps in BV epidemiology incited a question over OPV serological status of equids in Brazil. For this report, sera from 621 equids - representing different species, ages, sexes and locations of origin within Minas Gerais State, southeast Brazil – were examined for the presence of anti-*Orthopoxvirus* (OPV) antibodies. Only 74 of these were sampled during an Equine Vaccinia outbreak, meaning some of these specific animals presented typical lesions of OPV infections. The majority of sera, however, were sampled from animals without typical signs of OPV infection and during the absence of reported Bovine or Equine Vaccinia outbreaks. Results suggest the circulation of VACV among equids of southeast Brazil even prior to the time of the first VACV outbreak in 2008. There is a correlation of OPVs outbreaks among bovines and equids although many gaps remain to our understanding of its nature. The data obtained may even be carefully associated to recent discussion over OPVs history. Moreover, data is available to improve the knowledge and instigate new researches regarding OPVs circulation in Brazil and worldwide.

Keywords: *Vaccinia virus*, horsepox, bovine vaccinia, Equine Vaccinia, Horse diseases

OPEN ACCESS

Edited by:
Bernard La Scola,
Aix-Marseille Université, France

Reviewed by:
Filippo Turrini,
Vita-Salute San Raffaele University, Italy
Subir Sarker,
La Trobe University, Australia

****Correspondence:***
Iara A. Borges
borges2805@gmail.com

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

Received: *12 December 2017*
Accepted: *21 February 2018*
Published: *08 March 2018*

Citation:
Borges IA, Reynolds MG, McCollum AM, Figueiredo PO, Ambrosio LLD, Vieira FN, Costa GB, Matos ACD, de Andrade Almeida VM, Ferreira PCP, Lobato ZIP, dos Reis JKP, Kroon EG and Trindade GS (2018) Serological Evidence of Orthopoxvirus Circulation Among Equids, Southeast Brazil. Front. Microbiol. 9:402. doi: 10.3389/fmicb.2018.00402

INTRODUCTION

After 1980, following cessation of mass smallpox immunization, *Vaccinia virus* (VACV) – the *Orthopoxvirus* (OPV) used during the successful World Health Organization Smallpox Eradication Campaign (Fenner and Henderson, 1988) – emerged as a zoonosis in India, Pakistan and Brazil (Essbauer et al., 2010) VACV outbreaks involving dairy cattle and humans were first described in

Brazil in 1999 (Damaso et al., 2000). The number of Bovine Vaccinia (BV) reports continues to increase. BV commonly affects milking cows and cattle workers of small properties. Exanthemas are often located at the udder and teats of cows. All stages of the lesions – macule, papule, vesicle, pustule, ulcer, and scab – are highly contagious and direct or indirect contact with an abrasion or bare skin is enough to cause VACV infection in humans. Humans are often infected during manual milking; exanthemas are usually found on their hands and forearms (Damaso et al., 2000; Trindade et al., 2003).

Vaccinia virus is currently the only known OPV circulating in Brazil and has recently been detected in other South American countries such as Uruguay, Argentina and Colombia. However, other OPVs are known to occur in the Americas (Emerson et al., 2009; Gallardo-Romero et al., 2012). Recently, various species of rodents have been investigated as possible reservoirs for VACV and the virus has been detected in sylvatic and synantropic species (Abrahão et al., 2009, 2010a).

A VACV outbreak in horses occurred at a breeding center at south Brazil in 2008 and constitutes the first national report of VACV infected horses (Brum et al., 2010; Campos et al., 2011). Surprisingly, no relationship or contact to infected bovines was identified and the source of infection is not known up to date. A second and similar outbreak occurred in southeast Brazil during 2011 (Matos et al., 2013). Horses from different properties developed exanthemas on their muzzles and no infected bovine, human or other animal were determined as the possible source of infection, nor were any non-equine species observed to be infected as a result of the outbreak (Matos et al., 2013). The third outbreak reported the occurrence of oral lesions in donkeys and mules from the northeastern region of the country. Molecular findings indicated a VACV from Group 1 as the etiological agent and again no other species but *Equus sp*. have been connected to this outbreak (Abrahão et al., 2017).

Brazil has the largest herd of horses in Latin America and the third in the world. Together with the mules and donkeys are 8 million head, moving R$ 7.3 billion, only with the production of horses. Brazil is the eighth largest exporter of equine meat but when it comes to the export of live horses, the numbers are significant and expanded by 524% between 1997 and 2009. The largest Brazilian population of horses is in the Southeast, followed by the Northeast, Midwest, South and North, with one of its main functions being the daily work in agricultural activities, where about five million animals are primarily used for the management of cattle (MAPA).

To better comprehend the relationship between VACV and equids in Brazil, serum from more than six hundred equids from all mesoregions of Minas Gerais (MG) state, southeast Brazil, were analyzed for evidence of OPV exposure. Previous exposure of equids to OPV was evaluated by two different assays: the plaque reduction neutralization test (PRNT) to identify neutralizing antibodies (gold standard), and an enzyme-linked immunosorbent assay (ELISA). This latter method allows for assessment of non- neutralizing anti-OPV IgG antibodies.

This study provides an evaluation of serological evidence of VACV exposure in equids, expanding the epidemiological hypothesis of VACV circulation in Brazil, and potentially in other countries where VACV has been reported but its natural cycle is not well understood. The introduction of VACV into the Americas and its circulation among European horses prior to the 19th century have gained greater attention recently and highlighted the importance of horses in OPVs history (Damaso, 2017; Esparza et al., 2017; Schrick et al., 2017). Insights of the possible correlation of OPVs outbreaks among cows and equids have been evaluated and more data is currently available to improve the knowledge and also the development of control and prevention methods of OPVs circulation worldwide, specially VACV-like viruses.

MATERIALS AND METHODS

Samples

A total of 621 sera were examined from seven mesoregions of MG State. The mesoregions are defined by the animal defense bureau of the state of MG. **Figure 1** describes the abbreviations for regions that are used throughout this report. Four hundred seventy-eight sera from equids (*Equus caballus*, *E. asinus* and hybrids) were collected from numerous locations around MG state during the second half of 2003 and first half of 2004 (**Figure 1**). An additional 74 sera of *E. caballus* from RIV (**Figure 1**) were collected in July 2011, from the second known episode of Equid Vaccinia (EV) in Brazil. The serum specimens were collected approximately 1 month after the onset of symptoms, and clinical diagnoses were confirmed using molecular assays (Matos et al., 2013). The last 69 sera were randomly collected in RII during the second half of 2012 (**Figure 1**). All specimens were donated by the "Departamento de Medicina Veterinaria Preventiva, Escola de Veterinaria, Universidade Federal de Minas Gerais". All clinical specimens were derived from domestic equids on private properties and were collected by a veterinarian according to standard sanitary protocols in agreement to the requirements of national and local livestock agencies, "Ministerio da Agricultura, Pecuaria e Abastecimento" and "Instituto Mineiro de Agropecuaria", respectively. The sampling procedure was submitted and approved by the "Comite de Etica em Experimenação Animal (CETEA)" in accordance to the requirements of animal research of Universidade Federal de Minas Gerais (UFMG), MG state, Brazil (protocol number 131/2010, approval date 8/2010 and expiration date 8/2015).

Sterile vacuum blood collection tubes and needles were used to collect approximately 5 mL of blood from each animal while the animal was restrained with a bridle. The area of the jugular vein was cleaned with cotton soaked in 70% alcohol, and samples were collected from the jugular vein. A sterile and dry piece of cotton was used to apply pressure to the sampled area.

The name of the responsible employee and/or owner and the addresses of all properties were collected. Permission for sampling the horses was verbally granted. There was a verbal agreement to maintain confidentiality of the names and specific geographic location of each property. Coordinates from the regions further mentioned in this study are as follows (Datum: WGS84): RI or TRIANGULO

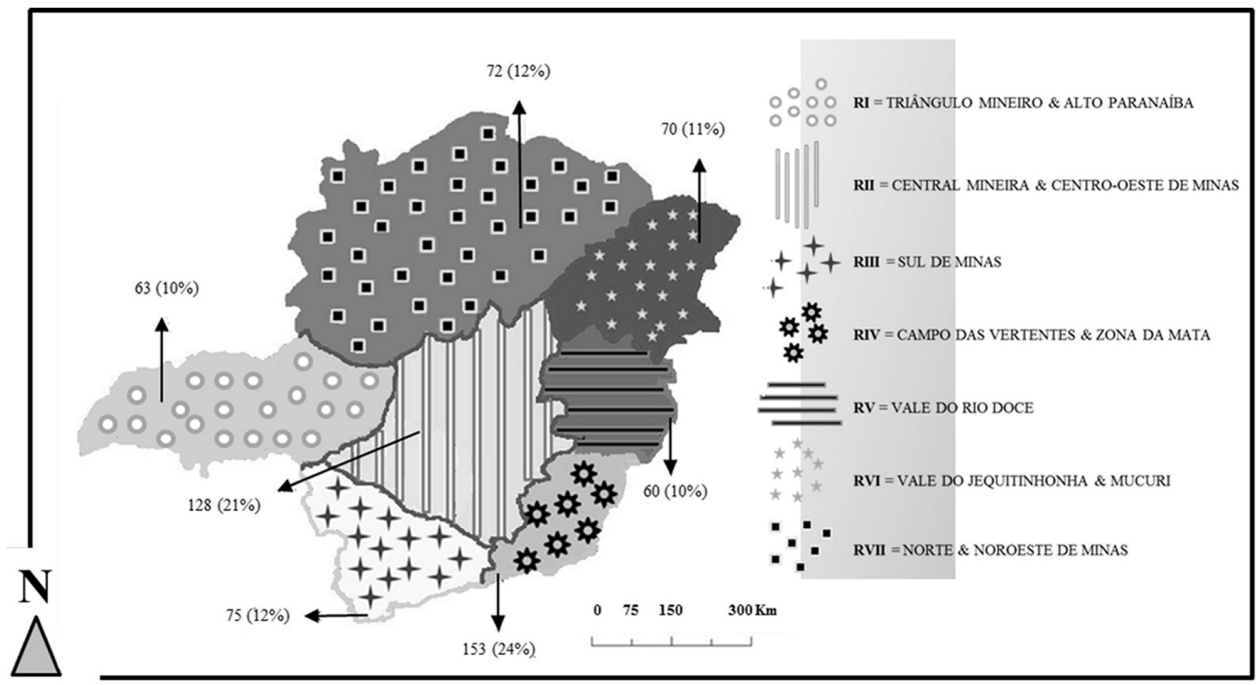


FIGURE 1 | Mesoregions of Minas Gerais State and equine specimens collected. The absolute number and correspondent percentage of serological specimens [n(%)] from each mesoregion is indicated by the arrows.

MINEIRO AND ALTO PARANAIBA 19°16′18.45″South and 48°20′59.95″West; RII or CENTRAL MINEIRA AND CENTRO-OESTE DE MINAS 18°51′15.52″South and 44°29′22.88″West; RIII or SUL DE MINAS 21°19′18.57″South and 45°48′09.67″West; RIV or CAMPO DAS VERTENTES AND ZONA DA MATA 21°30′01.91″South AND 43°27′57.24″West; RV or VALE DO RIO DOCE 19°28′22.44″South AND 41°46′46.62″ West; RVI or VALE DO JEQUITINHONHA AND MUCURI 16°58′31.86"South and 41°15′41.43″West; RVII or NORTE E NOROESTE DE MINAS 16°17′40.87″South and 44°41′03.21″West.

Epidemiologic and Demographic Variables

Four hundred seventy-eight sera were collected for multiple research purposes and there is no record of clinical signs suggestive of OPV acute infection among their database, as well as for the 69 sera randomly collected in RII. All these 547 sera have been submitted to PCR to OPV vgf gene according to Abrahão et al. (2010b) to detect a possible silent DNAmia; none trialed positive (data not shown). The geographic location, species, age, and sex of each equid sampled are also recorded.

The 74 sera (from the second EV outbreak notified in Brazil) were collected from equids with and without OPV-like exanthemas as described by Matos et al. (2013). According to them, viruses were isolated from sampled lesions and DNAmia was detected among several horses approximately 1 month after the first infected equid was noticed. Authors demonstrated through molecular methods a VACV sample to be the etiologic agent of this outbreak. The geographic location, species and sex of each equid sampled are recorded.

The location data for equids was aggregated by state subdivision. Subdivisions (mesoregions) are shown in **Figure 1** along with the absolute number and corresponding percentage of equids evaluated from each area. Since BV has traditionally been associated with dairy production in Brazil, it is helpful to consider dairy production activities in MG in relation to equid sampling locations. Milk producing properties are found throughout MG but several distinctions are found between the different mesoregions. Small dairy properties are concentrated at RIV and RII, where only one municipality had properties with areas larger than or equal to 69,106 Km^2. RIV is the area where the second known case of Brazilian EV occurred (Matos et al., 2013). Although traditionally known for its pure breed beef cattle herds, RI produced between 9 and 100 million of liters of milk during the year of 2006, the highest values found in MG State during the last census. RI properties also occupy the largest areas related to all other mesoregions as nearly 50% of its municipalities have properties with areas[1] over 69,106 Km^2.

Serological Assays

Plaque Reduction Neutralization Test

The PRNT was performed according to (Newman et al., 2003). Briefly, six well plates with BSC-40 cell monolayers (ATCC® CRL-2761) were inoculated with a 2.5% serum solution plus

[1]http://www.ibge.gov.br/

150 PFU of VACV Western Reserve (WR) per well. Before infection, sera/WR solutions were incubated overnight at 37^{o}C. To maintain the viability of the virus control, fetal bovine serum (FBS) was added to this solution at the same concentration (2.5%). The cell control contained 2.5% FBS media only. After infection, 1 h of adsorption was followed by the addition of 2 ml of 1% FBS media per well and incubation of all monolayers at 37^{o}C and 5% CO_2 for approximately 48 h. After typical VACV-WR cytopathic effects were clearly observed, all monolayers were fixed with 3.7% formaldehyde and stained with 1% crystal violet.

All samples were tested in triplicate and the number of PFU in each well was enumerated. Positive sera (positive for neutralizing antibodies) were defined as those samples that had PFU below the 50% PFU of the viral control.

IgG ELISA

Flat-bottom 96 well plates (Nunc MaxiSorp®) were coated with inactivated and purified VACV- WR viral particles; these plates were treated with a solution of PBS plus tween 20 (0.05%) and 5% casein. Sera were diluted 100x in a solution of PBS plus tween 20 (0.05%) and 1% casein.

To detect antibodies, anti-horse IgG conjugated to horse radish peroxidase (HRP) (SIGMA-ALDRICH®) and tetramethylbenzidine (TMB, BD®) were used for colorimetric detection at the concentration of 1:20000. All samples were tested in duplicates, and the results were read at 450 nm. Five known negative equid control sera were used. The average of all five negative controls minus three times their standard deviation determined the cutoff value for each plate.

Additionally, a known positive equid serum was tested for each plate as a positive control.

Statistical Analysis

Epidemiologic and demographic variables as well as qualitative laboratory findings were analyzed using parametric and non-parametric statistical tests. Pearson Chi-square test was employed for analyzing multiple-category non-parametric data. The Mantel-Haenszel common odds ratio (OR) and Fisher's exact tests (2-tailed) were used for categorical variables. A p-value < 0.050 was used to assign significance of association. All analyses were performed using IBM® SPSS Statistics version 19.

RESULTS

A total of 128 equids (20.6%) were seropositive for OPV by either the ELISA or PRNT assay (**Table 1**). Seropositive animals were found in all mesoregions of MG state. There was a lower frequency of seropositive equines in RI (4.8%, western MG) and RVII (13.9%, northern MG) and a higher frequency in RIV (29.4%, southeast) compared to the other regions of MG. The results from RI and RIV significantly deviate (−16% and +9%, respectively) from that which would be expected by chance alone (Chi-square 20.5, df =6, 2-sided p = 0.002). Moreover, a statistically significant difference was observed between these two regions (**Figure 2A**).

TABLE 1 | OPV seropositivity of equids by mesoregion.

Mesoregions	OPV antibody positive*[n(%)]	Total
RI	3 (4.8)†	63
RII	27 (21.1)	128
RIII	19 (25.3)	75
RIV	45 (29.4)†	153
RV	12 (20)	60
RVI	12 (17.1)	70
RVII	10 (13.9)	72
Total	**128 (20.6)**	**621**

**positive by either PRNT or ELISA. † > 8% deviation from that expected by chance alone.*

The location of the equids was significantly associated with seropositivity (Chi-square 20.5, df = 6, 2-sided p = 0.002). No significant relationship was observed for species, age or sex. The existence of seropositive equids throughout MG State, represents a unique result, since the possibility of VACV natural circulation among horses had been hypothesized, but not demonstrated.

Except in RII and RIV, a history of prior BV or EV occurrence was determined by property owner report only; such reports could not be independently confirmed. In cases where the property owner was not aware of the disease, pictoral representations of typical VACV lesions of humans, bovines and equids were provided. Regarding animals from properties whose owners confirmed BV or EV history – specific cases where samples were sent by IMA to official laboratories and results confirmed VACV infection – the frequency of seropositive equids from BV areas was almost two times higher (52.9%) than the frequency of seropositive equids in EV areas (27%). Furthermore, seropositive equids were found to be three times more likely to be from a property which formerly experiences a BV episode than those from areas where EV has occurred (OR = 3.0, 95% CI, 1.03–8.96, p = 0.044) (**Table 2** and **Figure 2B**).

DISCUSSION

These results demonstrate OPV exposure among equids in MG state, Brazil. As no OPV other than VACV has been proven to circulate in the country, results are hypothesized to be most likely related to the circulation of VACV. Moreover, the presence of serologically positive equids from 2003 to 2004 demonstrates that exposures, and perhaps unrecognized infections, occurred prior to the first described outbreak in Brazilian equids in 2008 (Brum et al., 2010). Results regarding the lower seropositivity observed in RI and higher seropositivity in RIV corroborate several studies in bovines, which associate intense VACV circulation with small property size (Kroon et al., 2011). Also, this result would be expected if there is a relationship between VACV circulation among equids and bovines. RIV is an important dairy region where several BV outbreaks have been reported (Lobato et al., 2005; Trindade et al., 2006; De Souza Trindade et al., 2007). The states that share a boundary with RIV – Espirito Santo and Rio de Janeiro – are also known for the frequent occurrence of BV

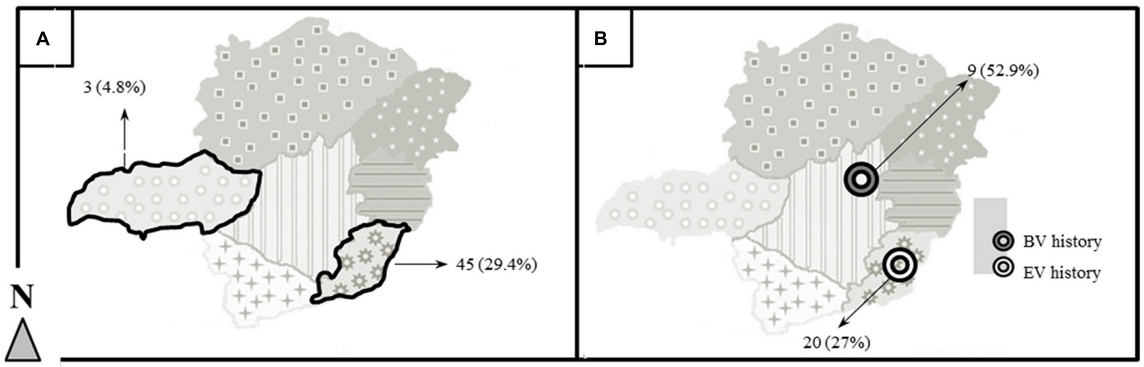


FIGURE 2 | Regional differences in seropositivity and a history of EV and BV outbreaks. **(A)** Two mesoregions of MG presented a discrepancy in equid seropositivity (RI and RIV). Both are highlighted; their absolute number and correspondent percentage of seropositive equids are indicated by arrows [n(%)]. **(B)** A difference of OPV seropositivity among equids derived from areas with EV or BV history was also observed. The white target indicates an area with a history of EV; the black target indicates an area with a history of BV. The absolute number and correspondent percentage of seropositive are indicated by arrows [n(%)].

TABLE 2 | OPV seropositivity of equids on properties with a history of EV or BV.

			OPV antibody positive*		
			Negative	**Positive**	**Total**
Epidemiological history	**EV history**	N	54	20	74
		%	73.0%	27.0%	100.0%
	BV history	N	8	9	17
		%	47.1%	52.9%	100.0%
Total		N	62	29	91
		%	68.1%	31.9%	100.0%

**positive by either PRNT or ELISA.*

(Kroon et al., 2011). RI, on the other hand, is known for its large herds of beef cattle[2] and fewer BV reports (Kroon et al., 2011).

When equids and bovines are maintained together on a single property, their degree of contact is often dictated by the size of the operation and type of production that occurs on the property. In small dairy herds, because of the nature of the work required, equid contact with bovines is considerably more intense than on larger, more sophisticated dairy farms. During dry seasons, these semi-intensive small properties require daily feed dispersal to support dairy production. The transportation of feed on these properties is usually accomplished by equids. Cattle are herded towards the corral and later back to the pastures, often by a human riding an equid. It is important to point out that despite its use in several other activities as sports, leisure and even therapy, one of its main functions, however, remains the daily work in agricultural activities, where about five million animals in Brazil are used primarily for the MAPA. In these situations, it is common for equids to freely access corral areas, share pastures, water, and feeding devices with bovines; therefore, equids housed on small properties with dairy cattle may indeed be more exposed to any pathogen circulating among the dairy cattle as compared to equids involved with beef cattle production or that are found on more technologically sophisticated farms.

The pure breed beef cattle that predominate in RI are highly valued economically[3]. This additional capital value allows for broader investments in a given property's dairy production. Therefore, the lower seropositivity observed among equids from this region may be a reflection of the investment-based technological innovations that have proliferated throughout the region. Increasing levels of sophistication on dairy properties may lead to the substitution of horses for tractors and crawlers thereby reducing bovine and equid contact. When equids from areas with a previous occurrence of BV versus a previous occurrence of EV were evaluated, a higher number of seropositive equids were found to be from BV properties.

Evaluating the typical systems in which horses are kept, opportunities for disease acquisition and transmission are somewhat different from those typical for small dairy cattle properties. Properties which dedicate part of their labor force to breeding horses tend to have individual stalls for their stallions; pastures present a density of equids considerably lower than those used with dairy cattle. Even if VACV replication and shedding is naturally efficient among horses and bovines, equids exposure would be significantly lower considering typical management practices for horses – those that are not working animals on a dairy farm.

[2]http://www.abcz.org.br/

[3]http://www.abcz.org.br/

The demonstration of VACV shedding on feces from deliberately infected cows (Rivetti et al., 2013) and the seroconversion of naïve animals after direct or indirect contact with this material (D'Anunciação et al., 2012) corroborate the possible co-dependency of VACV circulation observed among equids and bovines, as equids exposed to an OPV positive herd and environment may induce these equids seroconversion. It has been suggested that VACV is more successful in replicating and being shed by bovines as compared to equids; the experimental infection in equids with Brazilian samples of VACV indicated a low level of replication at the inoculation site with mild cutaneous lesions when compared with the course of infection of other hosts. The authors hypothesized equids have a low potential for viral maintenance and transmission to other species, albeit being susceptible to VACV infection (Barbosa et al., 2016).

Damaso, Esparza, Schrick, and collegues have revisited Edward Jenner' Inquiry in 2017 and vaccines tested by the doctor with specimens obtained from horses have come to all attention. Jenner attested "horse material" would most likely protect humans against smallpox only if previously inoculated in cows, as it rarely produced the "take" (exanthema developed at the site of inoculation) when directly sampled from horses and promptly inoculated in humans (reviewed by Baxby, 1999). Considering the possibility of VACV-like samples circulation in 19th Europe and their unknown use by Jenner, these observations may relate to what Barbosa called "a low level of replication" of Brazilian VACV samples in deliberately infected equids. The absence of human cases associated to all EV outbreaks reported in Brazil so far, also follows the pattern previously proposed and instigate the conclusion of a probable route of VACV from bovines to equids, as these last apparently are less effective in shedding the virus. Many characteristics of the infection of equids by VACV, however, remain to be investigated and assumptions should not be made based only at these studies.

Most importantly the evidence of silent (or possibly unreported) VACV exposure and disease in Brazilian equids in southeast Brazil has been demonstrated. VACV multiplication in equids occurs and even if it is less effective than in bovines, virus shedding into the environment constitutes a potential source of infection to other animals, domestic or wild, and to humans. Little is known about possible VACV reservoirs in Brazil as well as the exact importance of equids for VACV maintenance and circulation. Care must be taken in order to avoid VACV dissemination; veterinarians, horse and cattle caretakers, and all others involved in the equid industry should be informed of the risks related to EV.

Little investment has been made so far to prevent and control VACV infections in Brazil, either associated to BV or to the relatively recent EV. Disease prevention efforts are typically intensified only after a county suffers a significant outbreak. Municipalities that are potentially at risk remain uninformed about the disease until the impacts are apparent, including reduced milk production and painful human infections during BV outbreaks, bovine or equine morbidity during BV and EV outbreaks, respectively. Livestock disease surveillance programs have a role in outbreaks with suspected VACV cases, concomitant notifications to human public health authorities would speed the rapid inception of control measures, but research constitutes still the major investment related to VACV in Brazil. The continuation of serological monitoring and complementary research of VACV in equids is therefore fundamental to assess the spectrum of VACV impacts throughout the country, as well as to the development of prevention and control methods for its circulation nationwide.

AUTHOR CONTRIBUTIONS

IB did the standardization of the laboratorial techniques, field and laboratory work, data compilation, and paper writing. MR and AMM were responsible for guidance, statistical analysis, and paper writing. POF and FV did the field and laboratory work. LA, GC, ACM, VAA, ZL, and JR did the field work. PCF, ZL, JR, and EK provided guidance. GT contributed to the idealization of the project, guidance, and paper writing.

FUNDING

This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico, Pro-Reitoria de Pesquisa da UFMG (PRPq-UFMG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and Ministério da Agricultura, Pecuária e Abastecimento (MAPA). IB was supported by a fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. GT, EK, ZL, and JR are researchers of Conselho Nacional de Desenvolvimento Científico e Tecnológico.

ACKNOWLEDGMENTS

We thank all colleagues from Laboratorio de Virus for excellent technical support. We also would like to thank Fernanda Gonçalves de Oliveira for sera preparation, Valdísio for driving and helping us during the expeditions, and the Instituto Mineiro de Agropecuária (IMA) for technical support.

REFERENCES

Abrahão, J. S., Drumond, B. P., Trindade Gde, S., da Silva-Fernandes, A. T., Ferreira, J. M., Alves, P. A., et al. (2010a). Rapid detection of *Orthopoxvirus* by semi-nested PCR directly from clinical specimens: a useful alternative for routine laboratories. *J. Med. Virol.* 82, 692–699. doi: 10.1002/jmv.21617

Abrahão, J. S., Guedes, M. I., Trindade, G. S., Fonseca, F. G., Campos, R. K., Mota, B. F., et al. (2009). One more piece in the VACV ecological puzzle: could peridomestic rodents be the link between wildlife and bovine vaccinia outbreaks in Brazil? *PLoS One* 4:e7428. doi: 10.1371/journal.pone.0007428

Abrahão, J. S., Silva-Fernandes, A. T., Lima, L. S., Campos, R. K., Guedes, M. I. M., Cota, M. M. G., et al. (2010b). Vaccinia virus infection in monkeys, Brazilian Amazon. *Emerg. Infect. Dis.* 16, 976–979. doi: 10.3201/eid1606.091187

Abrahão, J. S., Trindade, G. S., Pereira-Oliveira, G., Figueiredo, P. O., Costa, G., Franco-Luiz, A. P., et al. (2017). Detection of *Vaccinia virus* during an outbreak of exanthemous oral lesions in Brazilian equids. *Equine Vet. J.* 49, 221–224. doi: 10.1111/evj.12571

Barbosa, C. H. G., Sant'ana, F. J. F., Cargnelutti, J. F., Flores, E. F., Teixeira-Neto, A. R., De Santana, R. B., et al. (2016). Experimental infection of horses with Vaccinia virus. *Ciên. Rural* 46, 519–525. doi: 10.1590/0103-8478cr20150289

Baxby, D. (1999). Edward Jenner's inquiry; a bicentennial analysis. *Vaccine* 17, 301–307. doi: 10.1016/S0264-410X(98)00207-2

Brum, M. C., Anjos, B. L., Nogueira, C. E., Amaral, L. A., Weiblen, R., and Flores, E. F. (2010). An outbreak of orthopoxvirus-associated disease in horses in southern Brazil. *J. Vet. Diagn. Invest.* 22, 143–147. doi: 10.1177/104063871002200132

Campos, R. K., Brum, M. C., Nogueira, C. E., Drumond, B. P., Alves, P. A., Siqueira-Lima, L., et al. (2011). Assessing the variability of Brazilian *Vaccinia virus* isolates from a horse exanthematic lesion: coinfection with distinct viruses. *Arch. Virol.* 156, 275–283. doi: 10.1007/s00705-010-0857-z

Damaso, C. R. (2017). Revisiting Jenner's mysteries, the role of the Beaugency lymph in the evolutionary path of ancient smallpox vaccines. *Lancet Infect. Dis.* 18, e55–e63. doi: 10.1016/S1473-3099(17)30445-0

Damaso, C. R., Esposito, J. J., Condit, R. C., and Moussatché, N. (2000). An emergent poxvirus from humans and cattle in Rio de Janeiro State: Cantagalo virus may derive from Brazilian smallpox vaccine. *Virology* 277, 439–449. doi: 10.1006/viro.2000.0603

D'Anunciação, L., Guedes, M. I., Oliveira, T. L., Rehfeld, I., Bonjardim, C. A., Ferreira, P. P., et al. (2012). Filling one more gap: experimental evidence of horizontal transmission of Vaccinia virus between bovines and rodents. *Vector Borne Zoonotic Dis.* 12, 61–64. doi: 10.1089/vbz.2011.0671

De Souza Trindade, G., Drumond, B. P., Guedes, M. I., Leite, J. A., Mota, B. E., Campos, M. A., et al. (2007). Zoonotic Vaccinia virus infection in Brazil: clinical description and implications for health professionals. *J. Clin. Microbiol.* 45, 1370–1372. doi: 10.1128/JCM.00920-06

Emerson, G. L., Li, Y., Frace, M. A., Olsen-Rasmussen, M. A., Khristova, M. L., Govil, D., et al. (2009). The phylogenetics and ecology of the orthopoxviruses endemic to North America. *PLoS One* 4:e7666. doi: 10.1371/journal.pone.0007666

Esparza, J., Schrick, L., Damaso, C. R., and Nitsche, A. (2017). Equination (inoculation of horsepox): an early alternative to vaccination (inoculation of cowpox) and the potential role of horsepox virus in the origin of the smallpox vaccine. *Vaccine* 35, 7222–7230. doi: 10.1016/j.vaccine.2017.11.003

Essbauer, S., Pfeffer, M., and Meyer, H. (2010). Zoonotic poxviruses. *Vet. Microbiol.* 140, 229–236. doi: 10.1016/j.vetmic.2009.08.026

Fenner, F., and Henderson, D. (1988). *Smallpox and its Eradication*. Geneva: World Health Organization.

Gallardo-Romero, N. F., Drew, C. P., Weiss, S. L., Metcalfe, M. G., Nakazawa, Y. J., Smith, S. K., et al. (2012). The pox in the North American backyard: volepox virus pathogenesis in California mice (*Peromyscus californicus*). *PLoS One* 7:e43881. doi: 10.1371/journal.pone.0043881

Kroon, E. G., Mota, B. E., Abrahão, J. S., Fonseca, F. G., and Trindade, G. S. (2011). Zoonotic Brazilian Vaccinia virus: from field to therapy. *Antiviral Res.* 92, 150–163. doi: 10.1016/j.antiviral.2011.08.018

Lobato, Z. I. P., Trindade, G. S., Frois, M. C. M., Ribero, G. R. C., Dias, B. M., Teixeira, F. A., et al. (2005). Outbreak of exanthemal disease caused by Vaccinia virus in human and cattle in zona da mata region, minas Gerais. *Arq. Bras. Med. Vet. Zootec.* 57, 423–429. doi: 10.1590/S0102-09352005000400001

Matos, A. C. D., Guedes, M. I. M. C., Rehfeld, I. S., Rodrigues, N. F. S., Costa, A. G., Madureira, M. C., et al. (2013). "Outbreak of Vaccinia virus infection in horses in minas Gerais state, Brazil," in *Proceedings of the XXIV Congresso Brasileiro de Virologia e Viii Encontro de Virologia do Mercosul, 2013*, Vol. 18 (Porto Seguro: Virus Reviews & Research), 240–241.

Newman, F. K., Frey, S. E., Blevins, T. P., Mandava, M., Bonifacio, A. Jr., Yan, L., et al. (2003). Improved assay to detect neutralizing antibody following vaccination with diluted or undiluted vaccinia (Dryvax) vaccine. *J. Clin. Microbiol.* 41, 3154–3157. doi: 10.1128/JCM.41.7.3154-3157.2003

Rivetti, A. V. Jr., Guedes, M. I., Rehfeld, I. S., Oliveira, T. M. L., Matos, A. C. D., Abrahão, J. S., et al. (2013). Bovine vaccinia, a systemic infection: evidence of fecal shedding, viremia and detection in lymphoid organs. *Vet. Microbiol.* 162, 103–111. doi: 10.1016/j.vetmic.2012.09.005

Schrick, L., Damaso, C. R., Esparza, J., and Nitsche, A. (2017). An Early American smallpox vaccine based on horsepox. *N. Engl. J. Med.* 377, 1491–1492. doi: 10.1056/NEJMc1707600

Trindade, G. S., Fonseca, F. G., Marques, J. T., Nogueira, M. L., Mendes, L. C. N., Borges, A. S., et al. (2003). Araçatuba virus: a vaccinia-like virus associated with infection in humans and cattle. *Emerg. Infect. Dis.* 9, 155–160. doi: 10.3201/eid0902.020244

Trindade, G. S., Lobato, Z. I., Drumond, B. P., Leite, J. A., Trigueiro, R. C., Guedes, M. I., et al. (2006). Short report: isolation of two *Vaccinia virus* strains from a single bovine vaccinia outbreak in rural area from Brazil: implications on the emergence of zoonotic orthopoxviruses. *Am. J. Trop. Med. Hyg.* 75, 486–490.

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 November 2017
doi: 10.3389/fmicb.2017.02251

To Be or Not To Be T4: Evidence of a Complex Evolutionary Pathway of Head Structure and Assembly in Giant *Salmonella* Virus SPN3US

Bazla Ali[1], *Maxim I. Desmond*[1], *Sara A. Mallory*[1], *Andrea D. Benítez*[1†], *Larry J. Buckley*[1], *Susan T. Weintraub*[2], *Michael V. Osier*[1], *Lindsay W. Black*[3] *and Julie A. Thomas*[1*]

[1] Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY, United States, [2] Biochemistry, University of Texas Health Science Center at San Antonio, San Antonio, TX, United States, [3] University of Maryland School of Medicine, Baltimore, MD, United States

OPEN ACCESS

Edited by:
Jonatas Abrahao, Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Juliana Cortines, Universidade Federal do Rio de Janeiro (UFRJ), Brazil
Gabriel Almeida, University of Jyväskylä, Finland

***Correspondence:**
Julie A. Thomas
jatsbi@rit.edu

†Present Address:
Andrea D. Benítez, Instituto Nacional de Investigación en Salud Pública, Centro Nacional de Referencia e Investigación en Vectores, Quito, Ecuador

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

Received: *18 August 2017*
Accepted: *31 October 2017*
Published: *15 November 2017*

Citation:
Ali B, Desmond MI, Mallory SA, Benítez AD, Buckley LJ, Weintraub ST, Osier MV, Black LW and Thomas JA (2017) To Be or Not To Be T4: Evidence of a Complex Evolutionary Pathway of Head Structure and Assembly in Giant Salmonella Virus SPN3US. Front. Microbiol. 8:2251. doi: 10.3389/fmicb.2017.02251

Giant *Salmonella* phage SPN3US has a 240-kb dsDNA genome and a large complex virion composed of many proteins for which the functions of most are undefined. We recently determined that SPN3US shares a core set of genes with related giant phages and sequenced and characterized 18 amber mutants to facilitate its use as a genetic model system. Notably, SPN3US and related giant phages contain a bolus of ejection proteins within their heads, including a multi-subunit virion RNA polymerase (vRNAP), that enter the host cell with the DNA during infection. In this study, we characterized the SPN3US virion using mass spectrometry to gain insight into its head composition and the features that its head shares with those of related giant phages and with T4 phage. SPN3US has only homologs to the T4 proteins critical for prohead shell formation, the portal and major capsid proteins, as well as to the major enzymes essential for head maturation, the prohead protease and large terminase subunit. Eight of ∼50 SPN3US head proteins were found to undergo proteolytic processing at a cleavage motif by the prohead protease gp245. Gp245 undergoes auto-cleavage of its *C*-terminus, suggesting this is a conserved activation and/or maturation feature of related phage proteases. Analyses of essential head gene mutants showed that the five subunits of the vRNAP must be assembled for any subunit to be incorporated into the prohead, although the assembled vRNAP must then undergo subsequent major conformational rearrangements in the DNA packed capsid to allow ejection through the ∼30 Å diameter tail tube for transcription from the injected DNA. In addition, ejection protein candidate gp243 was found to play a critical role in head assembly. Our analyses of the vRNAP and gp243 mutants highlighted an unexpected dichotomy in giant phage head maturation: while all analyzed giant phages have a homologous protease that processes major capsid and portal proteins, processing of ejection proteins is not always a stable/defining feature. Our identification in SPN3US, and related phages, of a diverged paralog to the prohead protease further hints toward a complicated evolutionary pathway for giant phage head structure and assembly.

Keywords: *Salmonella*, myovirus, giant phage, mass spectrometry, prohead protease, CTS (capsid targeting sequence), ejection protein, virion RNA polymerase (vRNAP)

INTRODUCTION

In recent years there has been a remarkable realization that standard phage isolation techniques were biased against larger phages and that "giant" or "jumbo" dsDNA tailed phages with genomes >200 kb can be readily isolated from a diversity of environmental samples and locales (Serwer et al., 2004, 2007, 2009; Krylov et al., 2007). The first giant phage genome published was that of *Pseudomonas aeruginosa* phage ϕKZ (280 kb) in 2002 (Mesyanzhinov et al., 2002). Since then even longer phage genomes have been reported, up to 480 kb (*Bacillus megaterium* phageG, Hendrix, 2009). Both ϕKZ and phageG were isolated decades prior to genome sequencing (Donelli, 1968; Krylov et al., 1984) and were considered rare oddities with fascinatingly complex virions but of little general relevance. However, ϕKZ-like phages have been incorporated into therapeutic mixtures of phages for phage therapy of *P. aeruginosa* infections (Krylov et al., 2007, 2012). There has been a surge in interest in phage therapy due to the problems of multi-drug resistant bacteria (Harper and Morales, 2012). Consequently, more than thirty phages related to ϕKZ infecting a range of hosts including *Salmonella* (e.g., SPN3US), *Erwinia amylovora* (e.g., Ea35-70), *Cronobacter sakazakii* (e.g., CR5) and *Vibrio* spp. (e.g., JM-2012) have now been isolated with the goal of developing novel phage-based therapeutics (Lee et al., 2011, 2016; Jang et al., 2013; Meczker et al., 2014; Yagubi et al., 2014; Bhunchoth et al., 2016; Danis-Wlodarczyk et al., 2016). For example, PhiEaH2 is one of two phages in "Erwiphage," the first marketed bacteriophage-based pesticide against *E. amylovora* in Hungary (Meczker et al., 2014).

There is also great interest in giant phages related to ϕKZ as their long genomes, which range in length from 211 to 316 kb, have many genes without counterparts in other tailed phage taxa. Not surprisingly, these phages have numbers of highly unusual traits relative to other phage types, including virions amongst the most complex of known phages, and atypical mechanisms for replication within the bacterial cell, as highlighted in recent studies (Kraemer James et al., 2012; Ceyssens et al., 2014; Zehr Elena et al., 2014; De Smet et al., 2016; Leskinen et al., 2016; Van den Bossche et al., 2016; Chaikeeratisak et al., 2017). Studies on ϕKZ determined that its virion is unusually large with a T = 27 capsid (Fokine et al., 2005a, 2007) and an odd structure within its head—a large proteinaceous inner body (IB), around which the DNA is tightly wrapped (Krylov et al., 1984). We determined the three-dimensional reconstruction of the ϕKZ IB and estimate it is comprised composed of 15–20 MDa of protein (Wu et al., 2012). That is, essentially the IB represents a large bolus or mass of many different proteins that are ejected into the host cell along with the genome (Wu et al., 2012). As observed for ϕKZ, other relatives of ϕKZ also show "bubblegrams" under the cryo-electron microscope indicative of an IB/ejection protein mass (Thomas et al., 2008; Sokolova et al., 2014).

Mass spectral analyses of phages ϕKZ and relatives *P. aeruginosa* phage EL and *P. chlororaphis* phage 201ϕ2-1 revealed that the virions of these phages are comprised of 60–70 different proteins—twice the number of different proteins found in the virion of the model phage T4 (Fokine et al., 2005a; Thomas et al., 2008, 2012; Lecoutere et al., 2009; Sycheva et al., 2012). The heads alone of phages related to ϕKZ contain ∼50 different proteins and the IB/ejection protein bolus is likely the locale for many of these proteins (Thomas et al., 2012). We hypothesize that these structures/proteins are likely multi-functional having roles in assembly, stability and host-takeover (Black and Thomas, 2012; Thomas et al., 2012). Support for the latter is that all related giant phages have a multi-subunit virion RNA polymerase (vRNAP) which is packaged within their heads for ejection into the host cell for production of early gene transcripts (Thomas et al., 2008; Ceyssens et al., 2014). The vRNAP and a second multi-subunit non-virion RNAP (nvRNAP) are a hallmark feature of all giant phages related to ϕKZ and are highly unusual as they are comprised of β and β' subunits which themselves are split into 2–3 subunits. This is very different from the single subunit RNAPs of phages T7 and N4 (Kazmierczak et al., 2002; Sousa and Mukherjee, 2003; Gleghorn et al., 2008) and likely enables transcription of giant phage proteins to be completely, or almost completely, independent of the host transcriptional machinery (Ceyssens et al., 2014; Leskinen et al., 2016).

We determined that head proteins in 201ϕ2-1 and ϕKZ, including the major capsid protein (MCP) and ejection protein candidates, undergo proteolysis by a protease (Thomas et al., 2012; Thomas and Black, 2013). We identified the ϕKZ enzyme, gp175, using bioinformatics analyses and then demonstrated that it cleaved two inner body proteins, making it the first phage protease to be expressed, highly purified and shown to be active *in vitro* (Thomas and Black, 2013). This enzyme is now the type protease for the MEROPs family S80 (Rawlings et al., 2012). Importantly, we identified that ϕKZ gp175 and its homologs in related giant phages, such as SPN3US gp245, share diverged sequence similarity and predicted structural elements with the prohead protease of T4 phage, gp21 (Thomas et al., 2012; Thomas and Black, 2013). Remarkably, the T4 prohead protease has been shown to conserve the structural elements and catalytic residues of Herpesvirus protease (Cheng et al., 2004; Liu and Mushegian, 2004; Rossmann et al., 2007). This finding, in conjunction with shared folds in MCP, portal and large terminase proteins, support a shared ancestor for Herpesviruses and tailed phages (Baker et al., 2005; Fokine et al., 2005b; Rixon and Schmid, 2014).

In T4, proteolytic processing by gp21 is an essential step in head assembly and results in a vast remodeling of the capsid architecture, setting the stage for genome packaging (Black et al., 1994; Miller et al., 2003). Briefly, T4 heads assemble via nucleation of the portal ring on the inner host membrane. Upon this ring a spherical mass of protein assembles, referred to as the core structure, which consists of the scaffold protein, gp22 (576 copies), internal proteins (∼1,000 copies), Alt (40 copies), gp67 (341 copies), gp68 (240 copies), and gp21 (100 copies) (Black et al., 1994). All of the T4 internal head proteins associate with the core structure via a capsid-targeting-sequence or CTS (Mullaney and Black, 1996). Upon completion of the core complex the shell then assembles around. The shell is composed of two proteins, the MCP, gp23 (960 copies) and gp24 (55 copies), which is located at the capsid vertices and is a paralog of gp24 (Fokine et al., 2006). At this point, the protease gp21 cleaves the core proteins gp22, gp67, and gp68 into short fragments, as well as *N*-terminal

propeptides from gp23, gp24, Alt and the IPs (Showe et al., 1976a,b). Fragments of the scaffold and core proteins, as well as the propeptides of the internal proteins exit the capsid and the capsid shell expands, increasing the internal volume of the capsid by 30–50% (Black and Rao, 2012). The prohead is then released from the host membrane and undergoes the final step of head assembly—DNA packaging where the terminase proteins packages the 170 kb genome into the head to remarkably high density (∼500 mg/ml) (Black and Rao, 2012; Black, 2015). The addition of the contractile tail assembled through a separate pathway and two decoration proteins, HOC and SOC, to the exterior of the capsid completes virion assembly.

These steps in T4 assembly, and many other features of this model phage, were elucidated with the aid of an elegant genetic system (Epstein et al., 1963, 2012). Not surprisingly, this system showed that the T4 genes associated with formation and maturation of the head, gps 17, 20, 21, 22, 23, gp24, 67, and 68, are essential (Miller et al., 2003). The proteins packaged within the dsDNA and ejected into the host cell, the IPs and the ADP-ribosyltransferase Alt, are not essential for assembly; however IPI is essential for the inhibition of a Type IV restriction endonuclease in *E. coli* strains containing the GmrSD enzyme gene (Black and Abremski, 1974; Abremski and Black, 1979; Bair et al., 2007; Rifat et al., 2008). Noting the power of such a genetic system and that in all phages related to ϕKZ there is a core set of conserved genes (including genes for major virion proteins and the vRNAPS), we sought to establish *Salmonella* phage SPN3US as a model giant phage genetic system (Thomas et al., 2016). We isolated, sequenced, and characterized 18 amber mutants of SPN3US, identifying 13 essential genes, only two of which, a SbcC and vRNAP β subunit, had been assigned putative functions previously. Illustrating the potential for this system, analyses of a putative neck gene mutant determined that ∼50 gene products are present in the mature SPN3US head. Additionally, analyses of the vRNAP β mutant facilitated identification of the previously unidentified *C*-terminal domain of the giant phage vRNAP β′ subunit and suggested a new phenomenon in phage head assembly—the five-subunit vRNAP enzyme complex assembles prior to incorporation into the prohead.

We continue to develop our novel genetic resource with the goal of implementing it to resolve questions regarding giant phage biology. To assist this process, in this study we sought to characterize the wild-type SPN3US virion using mass spectrometry and determine if proteolysis of head proteins occurs, as observed in related phages, and if so, define which proteins are processed and to what degree. In doing so, we aimed to gain insight into the head features and maturation characteristics that are shared in related giant phages and those that are shared between giant phages and T4 phage. Our results show that while SPN3US shares a head maturation protease with diverged sequence similarity to that of T4 and other major head features of T4, there are substantial variations in the proteolysis step in head maturation in SPN3US relative to T4 and also to other giant phages. Unexpectedly, we only identified approximately half the number of different proteins that undergo proteolytic processing in SPN3US vs. the number found in previously analyzed giant phage heads, despite all phages having approximately the same number of overall head proteins and sharing homologous proteins. In addition, our analyses of four SPN3US head gene mutants identified one of the processed head proteins as a novel head ejection protein (gp47), and another (gp243) that is essential for the assembly of the two most abundant, processed head ejection proteins into the prohead. The analyses of two vRNAP subunit mutants provided further evidence to support that the vRNAP is incorporated into the prohead as a multimer.

The identification of a diverged paralog to the prohead protease in SPN3US, and other giant phages, that is truncated and presumably no longer active (i.e., a cryptic protease) suggests a complex evolutionary pathway for the head proteolysis maturation step.

MATERIALS AND METHODS

Propagation and Purification of Phages

Bacterial stocks of *E. coli* and *Salmonella enterica* Typhimurium strains TT9079 and TT6675 and phages SPN3US and T4 were propagated using LB media. SPN3US wild-type and mutant phages were propagated in overlays containing 0.34% agar at 30°C. Phage dilutions were prepared in SM buffer. SPN3US amber mutants were isolated via hydroxylamine mutagenesis as described previously (Thomas et al., 2016).

Amber mutant particles were purified after propagation on the non-permissive host (TT9079) in LB broth supplemented with 1 mM $CaCl_2$ and 1 mM $MgCl_2$ at 30°C. Typically, an overnight culture was diluted 1:100 and grown to an OD_{600} of 0.3–0.5 at which point phage was added (MOI of 10). Phages were allowed to adsorb for 15 min then cells were pelleted twice at low speed (5,000 rpm, 3 min, RT) and resuspended in fresh media. Infected cells were then incubated with shaking for 3 h 30°C, and then treated with lysozyme (1 mg/ml) at room temperature with gentle shaking for 30 min. Mutant particles were concentrated by differential centrifugation then purified by sequential CsCl step then overnight buoyant density gradient ultracentrifugation as described previously (Thomas et al., 2016).

SPN3US Amber Mutant Genome Sequencing

SPN3US wild-type and mutant phage DNAs were extracted from high titer stocks that had undergone differential centrifugation (titers typically 5×10^{11} pfu/ml) and purified using a phage DNA extraction kit (Norgen). Mutant phage genomes underwent genome sequencing at the University of Rochester Genomics Research Center on an Illumina MiSeq machine (2×250). Genomes were assembled and SNP reports generated using SeqMan NGen and SeqMan Pro, respectively (DNASTAR). The reference sequence used for alignments was the wild-type SPN3US genome GenBank accession JN641803 (Lee et al., 2011).

Mass Spectrometry

Samples were boiled for 10 min in SDS sample buffer prior to electrophoresis on Criterion XT MOPS 12% SDS-PAGE reducing gels (Bio-Rad) and subsequent protein visualization

by staining with Coomassie blue. Gel lanes were divided into slices (10 for the wild-type phage, six for the mutants). Efforts were made to avoid transecting visible stained bands. No replicates of samples were analyzed. After de-staining, proteins in the gel slices were reduced with TCEP [tris(2-carboxyethyl)phosphine hydrochloride] and then alkylated with iodoacetamide before digestion with trypsin (Promega). HPLC-electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) was accomplished on a Thermo Fisher LTQ Orbitrap Velos Pro mass spectrometer or a Thermo Fisher Orbitrap Fusion Lumos mass spectrometer. Mascot (Matrix Science; London, UK) was used to search the MS files against a locally generated SPN3US protein database that had been concatenated with the SwissProt database (2012_11_170320; version 51.6). Subset searching of the Mascot output by X! Tandem, determination of probabilities of peptide assignments and protein identifications, and cross correlation of the Mascot and X! Tandem identifications were accomplished by Scaffold (Proteome Software). MS data files were either processed individually or the files for an entire gel lane were combined via the "MudPIT" option.

Peptides generated by cleavage by a prohead protease were identified through use of database searching (Mascot and X! Tandem) using an enzyme specificity of "semi-trypsin" followed by visual inspection of the results in Scaffold (Proteome software). This process was necessary because of unknown cleavage specificity of the SPN3US protease. Our previous studies have found variations in protease specificity even between the homologous proteases of ϕKZ and 201ϕ2-1 (Thomas et al., 2010, 2012; Thomas and Black, 2013). The results for identified proteins, numbers of unique peptides, total spectra, and sequence coverage for each experiment were exported from Scaffold with the following quality filters: peptide, 95%; protein, 99.9%; minimum number of peptides, 3. Microsoft Excel was then used to generate spectrum count profiles, as described previously (Thomas et al., 2010). An estimate of the relative abundance of SPN3US virion proteins was calculated by dividing the total number of spectra assigned for each protein (spectral count, SC) in the MudPIT analyses by its molecular mass (SC/ M) as performed for phages 0305ϕ8-36 (Thomas et al., 2007), 201ϕ2-1 (Thomas et al., 2007, 2010), RIO-1 (Hardies et al., 2016), and ϕKZ (Thomas et al., 2012). Our results from other phages have demonstrated that SC/M provides a useful indicator of relative abundance of different virion proteins. That is, proteins with similar SC/M values are typically present in similar relative abundances in the virion. In addition, proteins with SC/M ≤ 1 are likely to be present in only few copies, or even less than one copy, per virion. The copy numbers of several SPN3US virion proteins are known; major capsid protein (gp75), 1560 copies; tail sheath (gp256), 264 copies (Alasdair Steven and Weimin Wu, personal communication); and portal protein (gp81), 12 copies.

There are limitations to estimations of protein abundance using SC/M; therefore we also used densitometry, as conducted previously for the complex myoviruses 0305ϕ8-36 (Thomas et al., 2007) and ϕKZ (Thomas et al., 2012) and the podovirus RIO-1 (Hardies et al., 2016), for additional confirmation of head proteins assigned by SC/molecular mass as highly abundant. For these analyses, as two stoichiometric controls, the major sheath and portal proteins, migrate to the same region of the gel as the broad band of the highly abundant MCP (**Figure 1**) we used results from SPN3US mutant 64_112(am27) propagated under non-permissive conditions (it has a tailless phenotype) to assess the abundance of high copy number head proteins (Thomas et al., 2016).

Cloning, Expression, and Purification of the SPN3US Prohead Protease

The full-length form of the SPN3US protease (gp245) gene was amplified from SPN3US DNA by PCR using the primers pHS245F (5′-GCGCCATGGAAAACTTGTCACTACGTTATA ACTGCGTGGC-3′) and pHS245R (5′- GCGTCTAGATTA CCAGCTCCTTACACCCATGCCCATTACC-3′). The gp245 gene was then cloned into the vector pHERD20T (Qiu et al., 2008) using the NcoI and XbaI sites and transformed into *E. coli* DH10B. Codons for a six-histidine tag were subsequently added to the 5' end of gp245 gene via site-directed mutagenesis (SDM) as described in Thomas and Black (2013). PfuUltra Hotstart DNA polymerase (Agilent) was used for the amplifications, and subsequent digestion by DpnI (NEB) was undertaken to remove any remaining template DNA. Plasmid DNA was purified using the QIAprep spin miniprep kit (Qiagen) and the construct verified by DNA sequencing using the pHERD sequencing primers (Qiu et al., 2008). The gp245 construct was propagated in LB broth containing 150 μg/ml of ampicillin until mid-log phase, and protein expression was induced by the addition of arabinose (1% final concentration) for 1 h at 37°C. Cells were pelleted at 6,000 rpm (Sorvall rotor SS34) for 10 min, then resuspended in lysis buffer (20 mM Tris-Cl [pH 7.5], 150 mM NaCl, 1 mM EDTA, and egg white lysozyme [0.3 mg/ml; Sigma]), for ∼1 h, 4°C. The lysate was then treated with DNase (40 U/ml; Roche) at 37°C for 20 min, centrifuged (10,000 g, 10 min), and the supernatant was mixed with HisPur nickel resin (Thermo Scientific) overnight at 4°C. Purification of gp245 was performed according to the nickel column manufacturer's instructions at room temperature. Washes were performed in buffer containing 20 mM Tris-Cl (pH 7.5) and 300 mM NaCl with increasing imidazole concentrations (final elution buffer contained 250 mM imidazole). The eluted gp245 was dialyzed against wash buffer containing no imidazole and electrophoresed on a 12% Bis-Tris SDS-PAGE gel (Novagen). The major gel band was excised and digested separately with trypsin and chymotrypsin prior to mass spectrometric analyses as described above.

Transmission Electron Microscopy

Purified SPN3US mutant particles were adsorbed to 400 mesh carbon-coated grids and negatively stained with uranyl acetate (1%). Samples were examined at 80.0 kV using a FEI Tecnai T12 transmission electron microscope at the University of Maryland Electron Microscopy Core Imaging Facility.

Bioinformatics Analyses

PSI-BLAST searches were performed on a local implementations of the NCBI BLAST suite (Altschul et al., 1997), as were analyses using the Sequence Analysis and Modeling System (SAM)

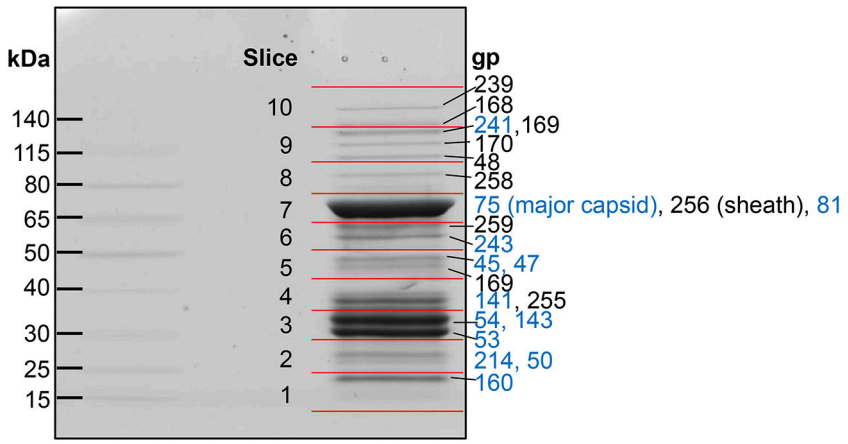


FIGURE 1 | SDS-PAGE gel of purified SPN3US virions. Individual gel slices that underwent mass spectrometric analyses are indicated by red lines. The most abundant SPN3US proteins identified in each gel slice are indicated with gene product (gp) numbers, with the most abundant protein in each slice, or gel band, listed first. Proteins that are components of the head structure are indicated (blue).

(Hughey and Krogh, 1996; Karplus et al., 1998) and HHpred (Söding, 2005). Phylogenetic trees were created from alignments created by MUSCLE (Edgar, 2004) using PAUP version 4.0 (Swofford, 2002) with Maximum Likelihood (50% majority rule) and the CDMut model. The majority rule consensus tree was created from 1,000 bootstrap runs.

RESULTS

Identification of SPN3US Virion Proteins by Mass Spectrometry

Eighty-six different SPN3US proteins were detected by mass spectrometry in CsCl step purified SPN3US, with a protein identification probability of 100% (**Figure 1**, Supplementary Table 1). This indicates that ∼33% of all SPN3US genes encode virion proteins; together, these proteins represent ∼46% because there are a number of long virion genes (e.g., gp168 is 5.2 kb). These SPN3US proteins detected in the virion ranged in molecular mass from 7 kDa for gp172 (a protein of unknown function) to 259 kDa for gp239 (the tail tape measure protein) (**Table 1**, Supplementary Table 1). The proteins were detected across a wide dynamic range of total spectra and sequence coverage, from four total spectra assigned for gp38 (a protein of unknown function) to 1,592 total spectra for gp75 (the MCP), and sequence coverage from 17% for gp49 to 94% for gp141 (Supplementary Table 1). Overall, fifty of the SPN3US proteins identified had sequence coverage of 50% or higher (Supplementary Table 1).

Spectral counting (SC) is an accepted semi-quantitative approach for estimation of relative protein abundance (Zybailov et al., 2005) and was used for this purpose for the SPN3US virion proteins. For each protein we calculated a SC/M value by dividing the total number of spectra assigned to each protein by its predicted molecular mass (**Table 1**, Supplementary Table 1). Several proteins were found to have undergone proteolytic processing (see below), and in these instances, an accordingly adjusted molecular mass was used to estimate relative abundance. After these adjustments, the protein in the virion with the highest SC/M value (22.6) was gp75, consistent with the expectation that the major capsid protein would be the most abundant protein in the virion. There are 1,560 copies of gp75 per virion based on the SPN3US capsid having the same triangulation number (T = 27) as ϕKZ (Alasdair Steven, Weimin Wu, personal communication). The SC/M values for gp53 and gp54 were 21.0 and 20.2, respectively, more than 2-fold the SC/M values obtained for the tail sheath and tube proteins. Two other SPN3US proteins, gp141 and gp160, had slightly higher SC/M values than the tail sheath SC/M value. The SPN3US tail sheath and tube proteins are expected to be present in ∼260 copies per particle based on the following considerations: 1. The SPN3US tail is of similar length as the ϕKZ tail, 2. SPN3US has sheath and tube proteins that are homologous to those of ϕKZ, 3. The ϕKZ tail is known to have 264 copies of each protein in its tail (Fokine et al., 2007).

The SC/M values for gps53 and 54 indicated that their respective copy numbers in the virion were higher than those for the tail sheath and tube proteins. This was unexpected because these proteins are incorporated inside the capsid shell with the DNA. However, the high abundance of gp53 and gp54 is consistent with the SDS-PAGE profile of this phage. To further explore the relative abundances of these two proteins, along with gp141 and gp160, we conducted densitometry analyses of a recently identified tailless mutant [*64_112*(am27); not shown]. These analyses supported our assignment of all four proteins as high abundance virion proteins and we conservatively estimate there are >600 copies each of gps 53 and 54 per capsid and >300 copies each of gps 160 and 141 per capsid.

Many of the SPN3US virion proteins are expected to be present in only a few copies per virion based on the expectation that their SC/molecular mass value is similar to that of gp81, the portal protein which is present in a dodecametric ring situated at a specialized vertex where the head joins the tail in all tailed phages. While some of the SPN3US low abundance proteins may

TABLE 1 | Abundant and processed virion proteins in purified SPN3US identified by mass spectrometry.

gp	Slice[a]	Mass, kDa (Proc. Mass)[b]	Total SC[c]	SC/M (Proc. M)[d]	Essential[e]	Paralog family	Comment (expected copies per virion)
75	7	83.9 (70.4)	1,592	18.97 (22.61)			**Major capsid** (1560 copies) Processed ATE-130.
53	3	45.2 (31.5)	662	14.65 (21.02)	Presumed	A	**Head ejection.** Processed AQE-125, AQE-95.
54	3	45.1 (31.9)	643	14.26 (20.16)		A	**Head ejection.** Processed AQE-124.
160	1	18.5	175	9.46		B	Head ejection?
141	4	32.6	306	9.39		B	Head ejection?
256	7	75.7	669	8.84	Presumed		**Tail sheath** (~264 copies)
255	4	32.7	209	6.39	Presumed		**Tail tube** (~264 copies)
MID-LOW ABUNDANCE PROTEINS							
243	6	54.6	251	4.60	Yes		Head ejection?
143	3	31.9	118	3.70		B	Head ejection?
45	5	50.3 (48.2)	144	2.86 (2.99)			Head. Processed at ASE-20
47[f]	5	62.8 (50.7)	110	1.75 (2.17)	Yes		Head ejection. Processed AVE-79, expected maturation cleavage ALE-111
50	2	39.4 (25.6)	56	1.42 (2.19)			Head. Processed ATE-127
169	9	149	210	1.41		C	Baseplate/ tail fiber candidate
241	9	159.1	189	1.19	Yes		Head ejection, **vRNAP β′N**
168	10	188.1	203	1.08	Yes	C	Baseplate/tail fiber candidate
218	2	25.2	27	1.07	Yes		Head ejection, **vRNAP βC**
42	5	49.3	52	1.05			Head ejection, **vRNAP β′M**
170	9	135.4	135	1.00		C	Baseplate/ tail fiber candidate
245	2	30.7 (23.4)	25	0.81(1.06)	Presumed		Head **Prohead protease**, Processed AQE-203,
81	7	100.2 (72.3)	93	0.93 (1.29)			Head **Portal (12 copies)** Processed ATE-161, expected AQE-254
240	6	59.6	33	0.55	Yes		Head ejection, **vRNAP β′N**
244	2	27	12	0.44	Yes		Head ejection, **vRNAP β′C**

*All highly abundant proteins are shown. Several mid- and low abundance proteins, such as the portal, prohead protease, and vRNAP subunits to highlight the range of abundances detected in the virion. All identified SPN3US proteins are listed in Supplementary Table 1. The slice in the SDS-PAGE gel (**Figure 1**) in which the spectral count for each protein peaked is provided. Relative abundance was determined from spectral counts adjusted for molecular mass (SC/M). Proteins identified as part of the head are indicated.*

*[a] The slice in the SDS-PAGE gel (**Figure 1**) in which the spectral count for each protein peaked.*

[b] "Proc. M" indicates molecular mass after processing by the prohead protease, gp245.

[c] "SC" represents the total number of spectral counts for each protein, as determined by the mass spectrometric MudPIT analyses.

[d] "SC/M" indicates total spectral count adjusted by molecular mass. Numbers provided in parentheses are the total spectral count adjusted by the processed molecular mass.

[e] "Essential"—indicates a protein encoded by a gene determined to be essential by the isolation and sequencing of an amber mutant of SPN3US.

[f] Mass spectral analyses re-assigned the start site of the gp47 gene to at nucleotide position 44887 in JN641803.1 which has additional four codons to the predicted start site. Processing sites of the prohead protease in gp47 are for the new peptide co-ordinates (see text).

not be true virion or assembly proteins, we expect that many are, based on the fact that the vRNAP subunits (gps 42, 218, 240, 241, and 244, **Table 1**) have low SC/molecular mass values, but are known to be essential and are packaged into the phage head. Similarly the low abundance of the prohead protease, gp245, is consistent with the expectation that only a few copies are present in the mature capsid, based on the estimation that the T4 protease is only present in about three copies per capsid (Black et al., 1994). The lack of detection of known non-virion proteins [such as the terminase protein (gp260) and DNA polymerase subunits (gp18 and gp44)] in the SPN3US virion provides further support that the low abundance proteins are true virion proteins.

Identification of Three Paralog Families in the SPN3US Virion

Twenty-five of the SPN3US proteins identified by mass spectrometry have homologs to other SPN3US virion proteins, as determined by PSI-BLAST. We categorized these paralogs into three families—Paralog families A, B and C (see examples in **Table 1**; all are included in Supplementary Table 1). Of these families, the three members of Paralog family C, the low abundance proteins gps 168, 169, and 170 (molecular masses of 188, 149, and 135 kDa, respectively), are the only ones for which a putative function has been deduced. These proteins are most likely related to the baseplate or fibers as they have similarity to ϕKZ gp131 (Sycheva et al., 2012). Supporting this expectation is the fact that these proteins were not identified in a tailless mutant [*64_112*(am27)] (Thomas et al., 2016). No putative function has been discovered, as yet, for Paralog families A and B.

SPN3US Paralog family A has two members, the highly abundant gp53 and gp54, both having similarity to pfam12699. Notably, the number of proteins with similarity to pfam12699 varies in different giant phage. In ϕKZ, for example, there are five homologs: the inner head proteins gps 93, 94, 95, 162, and 163 (Thomas et al., 2012) which are expected to be part of the IB and, therefore, excellent candidates for head ejection proteins. Interestingly, while ϕKZ gps 93, 95, and 162 are expected to be high abundance (>100 copies each per virion), their estimated

copy numbers are much lower than our estimates for SPN3US gp53 and gp54, implying that there is significant variation in the amounts of paralogs/homologs in different phages.

SPN3US Paralog family B is highly unusual because of the large number of members (20, gps83, 138–154, 161, and 237), all of which are expected to be head proteins since they were detected in the tailless mutant (Thomas et al., 2016). In addition, we believe that gp122 may be an additional member of this family based on PSI-BLAST searches; however, since gp122 was not detected in the mass spectral analysis of the wild-type phage, this assignment is speculative at this point. Unlike the other two SPN3US paralog families in which the members are present in similar relative abundances, the members of Paralog family B have relative abundances that range from low for most members, to middle (gp143) to high abundance (gps 141 and 160) (**Table 1**, Supplementary Table 1). There are no counterparts to any of the SPN3US paralog families in T4, although T4 has its own paralogs. As noted above, the T4 shell is composed of the MCP, gp23 and its paralog gp24, which forms the pentameric vertices of the capsid (Fokine et al., 2006).

Proteolytic Processing of SPN3US Head Proteins

Our analyses determined that eight SPN3US head proteins, gps 45, 47, 50, 53, 54, 75, 81, and 245, undergo proteolytic processing by the prohead protease gp245. Processing of these proteins was observed in the mass spectral analyses of the wild-type phage (except for gp245, see below), and additionally in mutant phage samples, both in this work and in our previous analyses (Hardies et al., 2016). In all instances, processing occurred *C*-terminal to a glutamic acid, at the motif A-X-E, where X is any amino acid (**Figure 2**). This is analogous to the cleavage motifs of the proteases of giant phages ϕKZ and 201ϕ2-1 (Thomas et al., 2010, 2012), as well of that of T4 phage (Black et al., 1994) (see **Figure 3**), all of which cleave *C*-terminal to a glutamic acid. The SPN3US protease processing sites were identified by detection of semi-tryptic peptides (e.g., gp75 and gp81 in **Figure 2**). Of the processed SPN3US proteins, only three have known functions: gp75, the MCP; gp81, the portal protein; and gp245, the prohead protease. The precursor form of gp75 has a molecular mass of 83.9 kDa, but after the removal of 130 residues by processing, the predicted molecular mass of this mature form is 70.4 kDa which is consistent with its SDS-PAGE gel migration (**Figure 1**). This *N*-terminal propeptide of SPN3US gp75 is twice the length of the 65 residue propeptide of the MCP gp23 of phage T4 (**Figure 3**).

The SPN3US portal gp81 (**Figure 2C**) has an immature form with a predicted molecular mass of 101 kDa from which a long *N*-terminal propeptide is removed (**Figure 2D**). This propeptide region of gp81 is at least 161 residues based on the identification of a small semi-tryptic fragment that was detected in the lowest molecular mass gel slice. However, based on the gel migration of the mature fragment of gp81 (where the peak of the spectral counts for gp81 were detected; slice 7, **Figure 2**) and its peptide coverage, we expect gp81 is also cleaved *C*-terminal to the sequence AQE-254 (**Figure 2C**). Processing at this site in gp81 would produce a mature form with a molecular mass of 72.3 kDa which is consistent with its gel migration. Despite the absence of similarity at the sequence level between the SPN3US portal and that of T4, as determined by PSI-BLAST, I-TASSER (Roy et al., 2010) structure prediction of the mature polypeptide of gp81 identified its most similar structural homolog as T4 gp20 (Sun et al., 2015) and predicted domains consistent with the crown, wing, stem and clip observed in the portal proteins of several other tailed phages (Orlova et al., 2003; Lhuillier et al., 2009). That is, it is the long propeptide of gp81 that markedly delineates the SPN3US portal from that of T4. Supporting this we confirmed for T4 gp20 for the first time biochemically by mass spectrometry that it is not processed (**Figure 3**). The function of the long propeptide of SPN3US gp81 can only be speculated upon, but we presume it has a role in head assembly, possibly helping to anchor ejection and/or core proteins during assembly.

Four other SPN3US head proteins have *N*-terminal propeptides longer than 100 residues, including gps 50, 53, and 54, with propeptides of 127, 125, and 124 residues, respectively. These propeptides are considerably longer than the 10–20-residue *N*-terminal propeptides of T4 head proteins, such as gp24 (**Figure 3**) and its IP proteins (Supplementary Figure 1), referred to as capsid targeting sequences (CTS) for their role in ensuring a protein is incorporated into the prohead (Mullaney and Black, 1996). Of the SPN3US head proteins that undergo *N*-terminal processing, only gp45 has a propeptide of comparable length (20 residues) to the T4 CTS sequences.

Gp47 also has a long *N*-terminal propeptide, and while it is processed at residue 79 (new co-ordinate, see below), because there was no MS sequence coverage until residue 119 it is more likely that its maturation cleavage is after the sequence which adheres to the cleavage motif: ALE-111 (new co-ordinate) (**Figure 4**). Our analyses also confirmed a new start site for gp47 as three amino acids were detected by mass spectrometry in the tryptic peptide SME*M*TGNAPHTK which are upstream of its predicted start methionine (underlined) of gp47 (**Figure 4A**). The codon immediately upstream of this peptide is a methionine leading us to conclude the start site of the gp47 gene is at nucleotide position 44887 in JN641803.1 rather than the predicted nucleotide position 44899. The *N*-terminal methionine of gp47 is likely cleaved by the host methionine aminopeptidase (Wingfield et al., 1989; Movva et al., 1990). The new start methionine codon for gp47 has a credible upstream ribosomal binding site that results in a small overlap of this region with the 3' end of the gp46 gene.

Several of the long propeptides of SPN3US proteins have additional sequences that are consistent with the protease processing motif and may also be cleaved by protease. For instance, the *N*-terminal 15 residues (MANFVKSKLARESVE) of the processed paralogs gp53 and gp54 are identical and contain a sequence (AXE-12) consistent with the known cleavage motif for SPN3US and a sequence (SXE-15) consistent with the ϕKZ protease cleavage motif. We infer that gp53 and gp54 are processed at one or both of these sites, because the ϕKZ homolog, gp93, is processed at SLE-13 (Thomas and Black, 2013). The advantage of additional cleavage sites within propeptide regions would be to produce smaller fragments more easily cleared from the capsid during maturation. That the N-termini of gp53 and

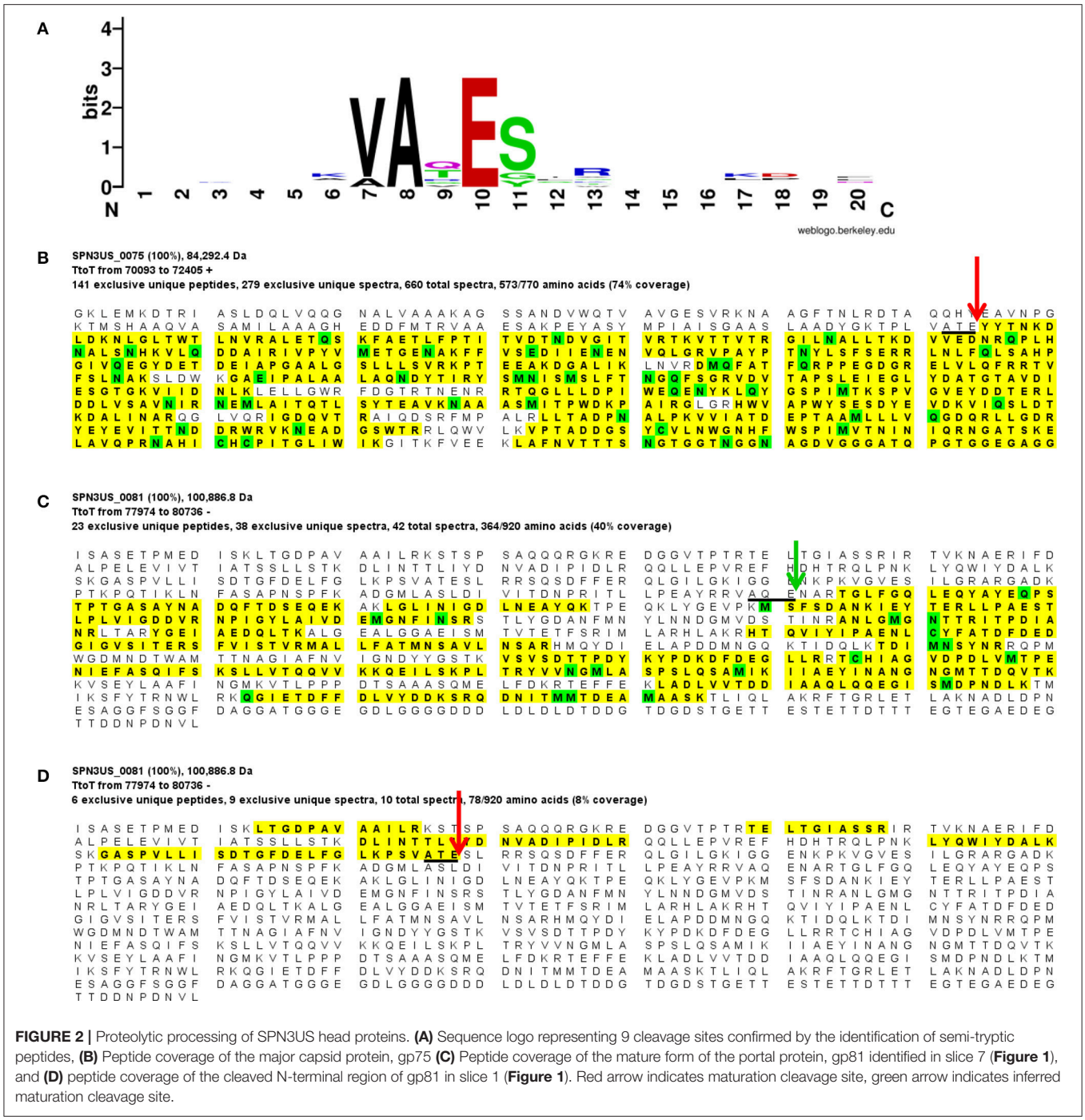


FIGURE 2 | Proteolytic processing of SPN3US head proteins. **(A)** Sequence logo representing 9 cleavage sites confirmed by the identification of semi-tryptic peptides, **(B)** Peptide coverage of the major capsid protein, gp75 **(C)** Peptide coverage of the mature form of the portal protein, gp81 identified in slice 7 (**Figure 1**), and **(D)** peptide coverage of the cleaved N-terminal region of gp81 in slice 1 (**Figure 1**). Red arrow indicates maturation cleavage site, green arrow indicates inferred maturation cleavage site.

gp54 are identical is notable as overall the proteins have diverged extensively from another at the sequence level, as evidenced by their having only 34% identity by BlastP. The sequence conservation at the *N*-termini of gp53 and 54 is reminiscent of the sequence conservation in several of the T4 CTSs and suggests they have a conserved/important role in assembly/maturation.

It is feasible that there may be a small number of additional low abundance proteins in SPN3US that undergo proteolytic processing that were not identified, as peptide coverage is naturally lower in lower abundance proteins. Peptide detection is also impacted by the cleavage specificity of trypsin and resultant peptide length. However, we believe the number of any additional processed proteins to be small based on our examination of eight mass spectrometric SPN3US samples and the identification of intact N- and/or C-termini in many proteins. Identification of any potential additional processed proteins would likely require biochemical assays of recombinant proteins

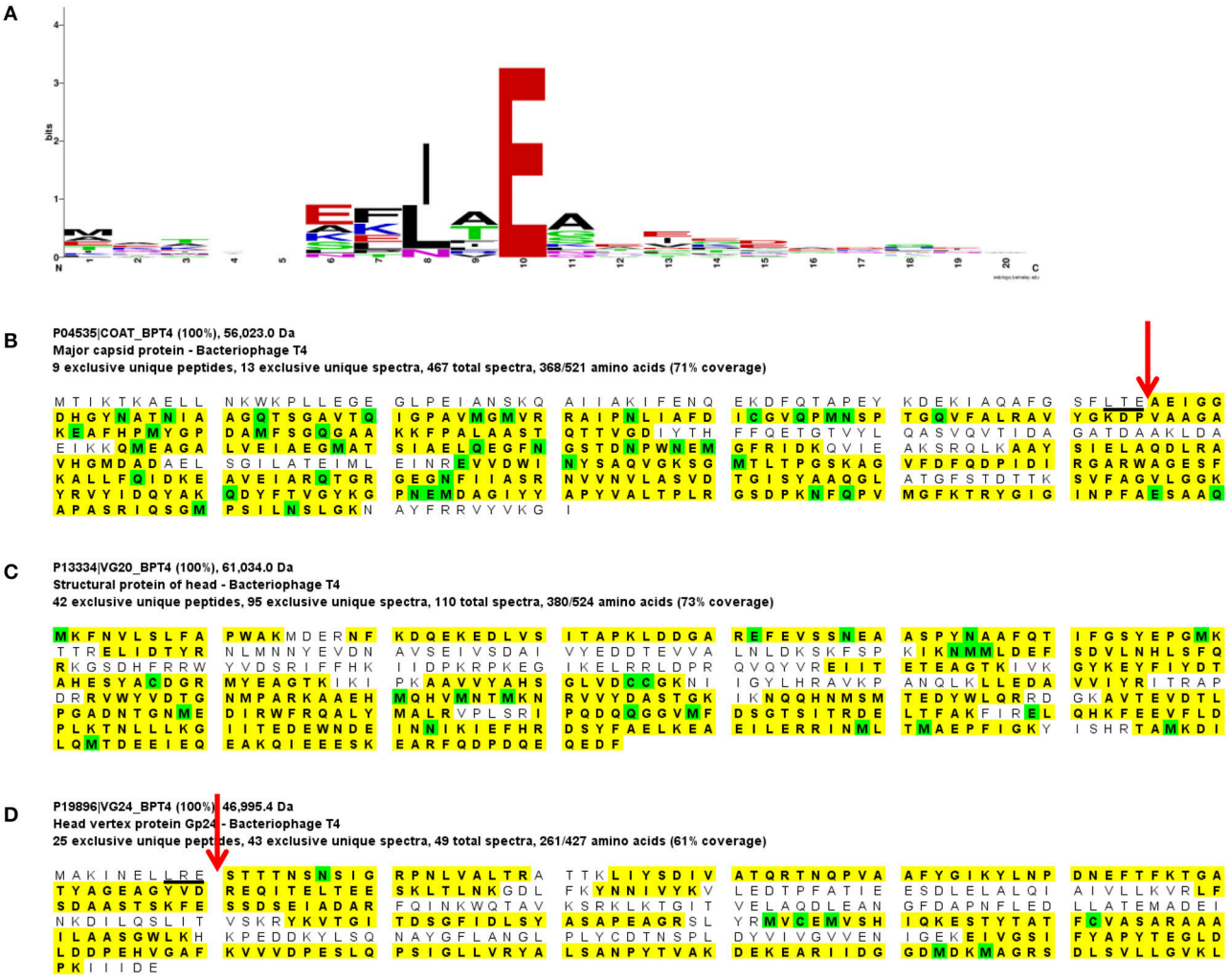


FIGURE 3 | Proteolytic processing of T4 head proteins. **(A)** Sequence logo representing 13 cleavage sites in 9 head proteins confirmed by *N*-terminal sequencing (Black et al., 1994) and/or semi-tryptic peptides (this study). **(B)** Peptide coverage of the major capsid protein, gp23, **(C)** Peptide coverage of the portal protein, gp20, which is not processed, and **(D)**. Peptide coverage of the capsid vertex protein, gp24. Red arrow indicates maturation cleavage site.

and/or a mass spectrometric analyses of a sample that underwent electrophoresis through a much longer SDS-PAGE gel and consequent division into many slices to enable clear identification of any aberrant gel migration relative to that expected based on the predicted molecular mass.

Auto-Proteolytic Processing of the SPN3US Prohead Protease

The SPN3US prohead protease, gp245 (263 residues), was identified in the wild-type phage particle as a low abundance protein (**Table 1**). The mass spectral sequence coverage for gp245 ended at residue K-190 (**Figure 5**) which is *N*-terminal to three possible cleavage motifs [AQE-246, ATE-234, and AQE-203], leading us to suspect that gp245 underwent C-terminal autocleavage, as we had observed for the ϕKZ protease (Thomas and Black, 2013). To test this hypothesis, we cloned the full-length gp245 gene with additional codons for an *N*-terminal 6-histidine tag in the expression vector pHERD20T and purified the recombinant enzyme on a nickel column. The migration of 6His-gp245 in SDS-PAGE (**Figure 5B**) was consistent with a lower molecular mass than the predicted 30.7 kDa of the full-length enzyme. This protein band was excised from the gel, digested with chymotrypsin and trypsin and analyzed by mass spectrometry. The resulting peptide coverage obtained for 6His-gp245 identified the *C*-terminus of gp245 (**Figure 5C**) at glutamate residue 203, confirming its autoprocessing. Cleavage at glutamate 203 produces a mature species with a molecular mass of 23.4 kDa, however it is feasible that one or both of the downstream glutamate residues (residues 234 or 246) are also cleaved. Conversely, we were able to deduce the *N*-terminus of gp245 does not undergo auto-cleavage based on the sequence coverage of gp245 in several samples [e.g., *21*8(am101), **Figure 5D**] and the fact that there are no sequences that contain the AXE-cleavage motif in this region.

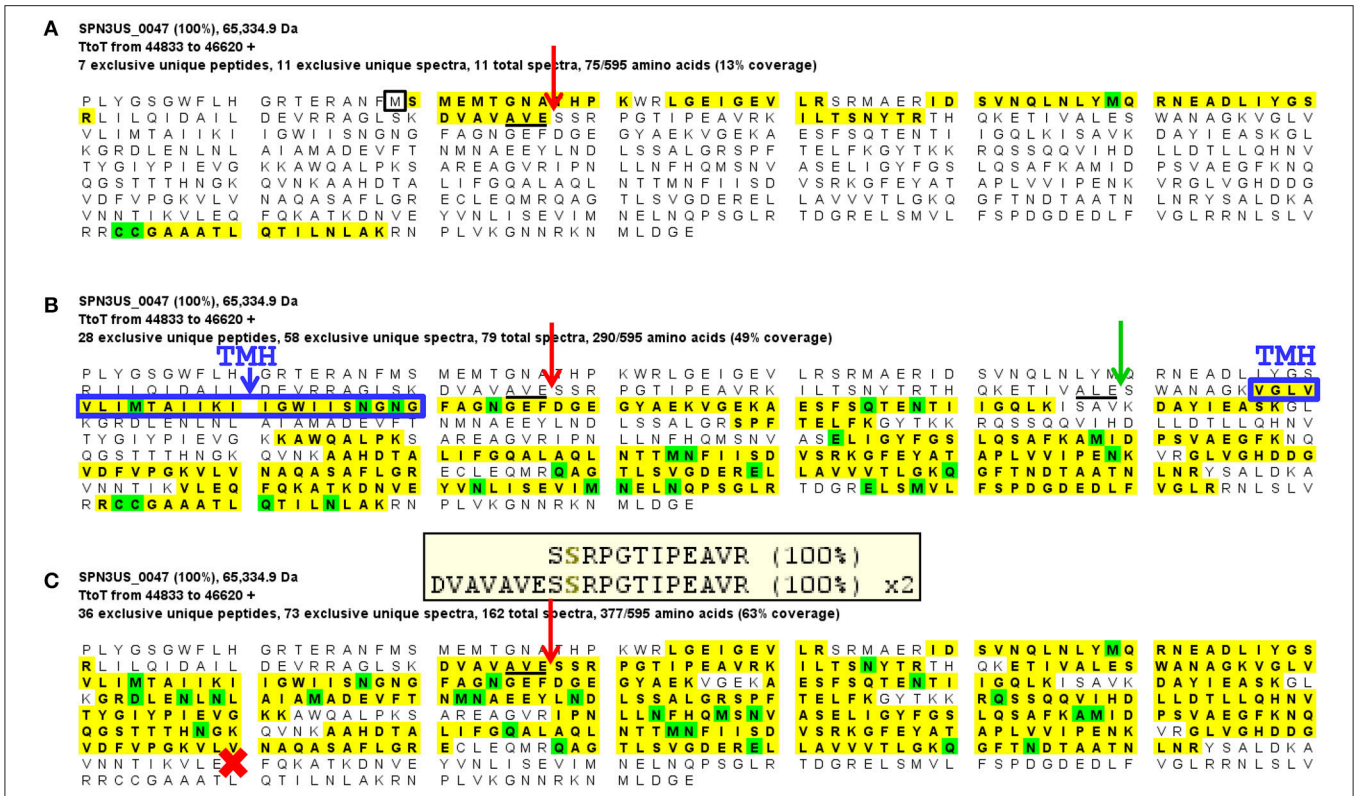


FIGURE 4 | Mass spectral identification of SPN3US ejection protein candidate, gp47. **(A)** Peptide coverage of the precursor propeptide of gp47 identified in the wild-type phage detected in a low molecular mass SDS-PAGE gel slice (slice 1, **Figure 1**). Red arrow indicates the prohead protease cleavage site (AVE-79) identified via a semi-tryptic peptide; black square indicates newly identified start methionine, **(B)** Peptide coverage of the mature polypeptide of gp47 identified in the wild-type phage detected in a higher molecular mass SDS-PAGE gel slice (slice 5, **Figure 1**). Green arrow indicates expected protease maturation cleavage site at ALE-111; blue box indicates predicted transmembrane helix (TMH), and **(C)** Peptide coverage of gp47 in the mutant *47*(am1) propagated under non-permissive conditions causing gp47 to be truncated at Q482 (red cross). Note the prohead protease cleavage site from **(A)** is only partially cleaved in this mutant as demonstrated by the identification of both a semi-tryptic peptide that illustrates processing and a tryptic peptide that illustrates no processing at the cleavage motif, 100% refers to the peptide identification probability.

We were unable to detect the T4 prohead protease gp21 in T4 heads by mass spectrometry, although it was estimated to be present in the capsid at approximately three copies (Black et al., 1994). The lack of identification may have been a consequence of variations in virus propagation, purification, or there may be a biological cause, such as if in our strain of T4 the protease is efficiently and completely cleared from the prohead. (When gp21 is first packaged into the head, it is estimated to be present in ∼100 copies, Black et al., 1994). However as a consequence of not detecting gp21, we were unable to determine if either or both of its suggested autocleavages [*C*-terminal (Keller and Bickle, 1986) or *N*-terminal (Fokine and Rossmann, 2016)] occur. Interestingly, in these studies we made a new observation regarding proteolytic processing of the T4 ejection protein Alt finding it to undergo both *N*-terminal and *C*-processing by gp21 to produce a mature protein of 68.2 kDa (see Supplementary Figure 1C). In addition to the known six residues removed from the *N*-terminus of Alt (by processing at the motif ITE-6), 63 residues are also removed from the *C*-terminus (by processing at the motif LTE-619). This new finding explains the previously noted aberrant SDS-PAGE migration of Alt (faster than expected) and raises questions as to the role(s) of the two propeptides flanking Alt—are they targeting, tethering, and/or controlling its activity? The processing of both termini of Alt highlights the need for future work to clarify the processing status of the T4 protease.

Characterization of SPN3US Essential Head Protein Mutants

To further characterize SPN3US head composition and assembly, we examined the effect on the virion when several essential head proteins (gps 47, 218, 243, and 244, **Table 1**) were knocked out. To do this we utilized amber mutant phages *47*(am1) (Thomas et al., 2016) and newly isolated mutants *218*(am101), *243*(am114), *244*(am84). The genome of each mutant was sequenced and an individual amber mutation in a single gene was identified in all, allowing us to conclude that each gene and its product are essential (**Table 2**, Supplementary Table 2). Putative functions could only be assigned to the products of two mutated genes, gp218 and gp244, which represent the *C*-terminal subunits of the vRNAP β and β′, respectively (Thomas et al., 2016). These vRNAP subunits were knocked

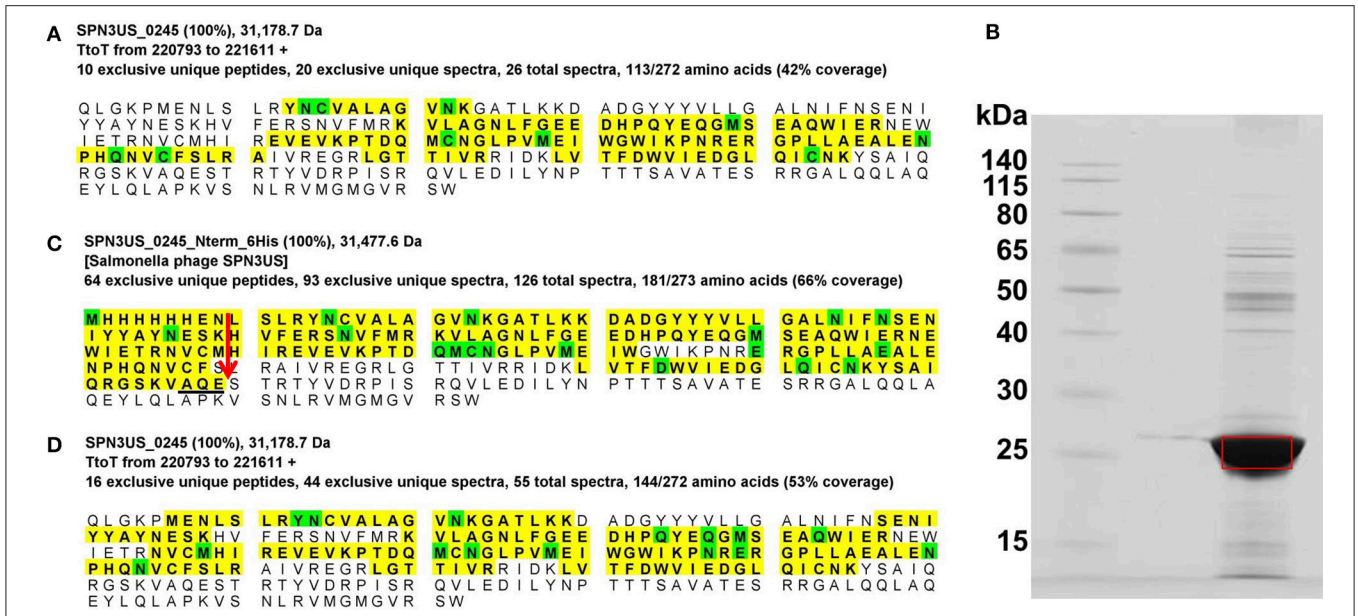


FIGURE 5 | Auto-proteolytic processing of the SPN3US prohead protease, gp245. **(A)** Peptide coverage of gp245 in purified virions (slice 2, **Figure 1**), **(B)** SDS-PAGE gel of gp245 expressed from the full length gene with an *N*-terminal 6-histidine tag, **(C)** Peptide coverage of recombinant gp245 with an *N*-terminal 6-Histidine tag, and **(D)** Peptide coverage of gp245 in purified virions of *218* that had been propagated under non-permissive conditions. Red arrow indicates the prohead protease cleavage site (AQE-203) identified by a semi-tryptic peptide.

TABLE 2 | SPN3US virion proteins not identified in mass spectrometric analyses of amber mutants grown under non-permissive conditions.

Mutant	Mutated gene product, gp	vRNAP subunits not identified, gp	Other SPN3US virion proteins not identified, gp	SPN3US proteins identified, that were not identified as virion proteins in the wild-type phage, gp
47(am1)	47		38 (unknown function) 157 (head protein of unknown function)	100 (thymidylate kinase) 122 (paralog Family B candidate) 158 (unknown function)
218(am101)	218	42 (β′M) 218 (βC) 240 (β′N) 241 (βN) 244 (β′C)	38 (unknown function) 41 (unknown function) 98 (unknown function) 157 (head protein of unknown function)	33 (unknown function) 100 (thymidylate kinase) 158 (unknown function)
244(am84)	244	42 (β′M) 218 (βC) 240 (β′N) 241 (βN) 244 (β′C)	38 (unknown function) 41 (unknown function) 98 (unknown function) 157 (head protein of unknown function)	28 (unknown function)

out by propagating the respective mutant under non-permissive conditions and the resulting particles purified via consecutive CsCl step and buoyant density ultracentrifugation gradients. For both mutants, seemingly intact virions were formed, as judged by the gross morphological features (e.g., DNA-filled head, sheath uncontracted, baseplate/tail fibers) and dimensions typical of the wild-type phage (e.g., **Figure 6A**). However, the purified *218*(am101) and *244*(am84) particles were not viable when propagated on either the permissive or non-permissive host. Mass spectrometric analysis of each mutant proteome revealed that neither mutant had the full complement of virion proteins that were identified in the wild-type phage (**Table 2**, Supplementary Table 3). Notably, none of the five vRNAP subunits (gps 241, 218, 240, 42, and 244) were identified when either gp218 or gp244 were knocked out, providing further support for our recent identification of gp244 as the fifth subunit of this enzyme and our hypothesis that the vRNAP is assembled as a multimer prior to its incorporation into the prohead (Thomas et al., 2016).

SPN3US gp47 is a low abundance head protein that, as noted above, undergoes proteolytic processing to remove an *N*-terminal propeptide of 111 residues. Interestingly, despite its essential status, gp47 did not have an easily identifiable homolog in ϕKZ. However, gp47 clearly has counterparts in

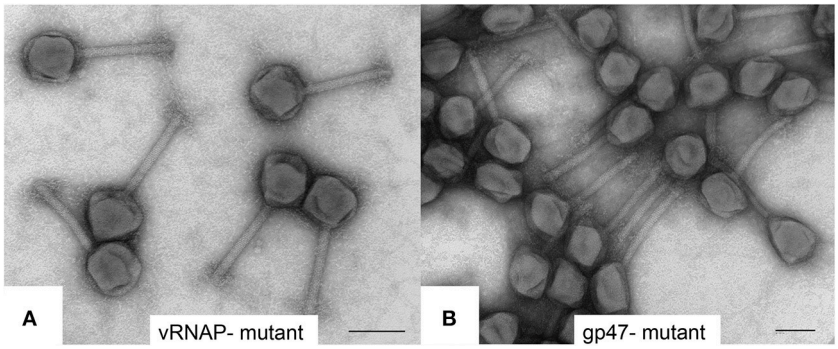


FIGURE 6 | Transmission electron microscopy of negatively stained SPN3US amber mutants **(A)** vRNAP minus mutant [*218*(am101)] and **(B)** ejection protein gp47 minus mutant [*47*(am1)] after propagation on nonpermissive *S. enterica* serovar Typhimurium (strain TT9079). Space bar represents 100 nm.

the more closely related *Erwinina* phages, such as PhiEaH2 gp231, and *Cronobacter* phage CR5 gp48 (57 and 27% identity by BlastP, respectively) (Thomas et al., 2016). This led us to examine the gp47 gene locale which revealed that it is nested in a region encoding several virion protein genes which have identifiable homologs in related giant phages, all with a complex, but apparently conserved transcriptional orientations (**Figure 7**). Examination of this syntenous cluster indicates that the gp47 gene is in the equivalent location to that of ϕKZ gp86 and its homologs in related phages, 201ϕ2-1 gp148 and ϕPA3 gp89. Notably, the similarity between these three homologs is low, only 24–25% identity by BlastP, suggesting these genes are under selective pressures to evolve more rapidly than more highly conserved proteins, such as the terminase and major capsid proteins which have 60–69% identity by BlastP between ϕKZ, and 201ϕ2-1 and ϕPA3. Since the major capsid and terminase proteins of OBP and EL, the *Pseudomonas* phages most closely related to ϕKZ, have only 22–29% identity by BlastP to their ϕKZ homologs, we could speculate that OBP gp102 and EL gp54, whose genes are in the same locale as that of ϕKZ gp86, may have shared ancestry with gp86, but are now passed the horizon of search detection limits. Supporting this possibility is that OBP gp102 and EL gp54 have sequence homology (21% identity by BlastP).

Notably, ϕKZ gp86 is an inner head protein that undergoes processing at SQE-122 by the prohead protease gp175 and is expected to be ejected with the DNA into the host cell with its multi-subunit RNAP and other head ejection proteins (Thomas et al., 2012). Similarly, the 201ϕ2-1 homolog to ϕKZ gp86 (gp148) is processed at SQE-113 and is expected to be ejected into the host cell with its multi-subunit RNAP and other head ejection proteins (Thomas et al., 2010). In ϕKZ and 201ϕ2-1, the N-terminal fragments of gp86 and gp148 are completely cleared from the head, in contrast to the *N*-terminal fragment of gp47 of SPN3US which was detected in the lowest molecular mass gel slice. Additional experiments are needed to determine whether the *N*-terminal fragment that remains in the SPN3US head is of biological significance or if it is unable to exit the head.

Similar to the vRNAP mutants (am101), the gp47 mutant, *47*(am1), produced non-viable particles when propagated under non-permissive conditions although its virions appeared intact as judged by TEM (**Figure 6B**). Mass spectrometry analysis revealed that all proteins identified in the wild-type particles were present in *47*(am1), with the exception of gp172, a protein that was detected at low abundance in the wild-type phage. The lack of detection of gp172 in these mutants may be biologically relevant, or possibly it was not detected if the low molecular weight region of the gel was not included in the analyses. At 7 kDa, gp172 is the smallest SPN3US virion protein. Further experiments are required to resolve its status. Additionally, very low amounts of three proteins (gps 100, 122, and gp158) were identified in the *47*(am1) proteome that were not detected in the wild-type virion (**Table 2**). As the predicted function of gp100 is a thymidylate synthase, it seems unlikely that it and the other two proteins identified only in this mutant are true virion proteins. We suspect these proteins were non-specifically associated with the virion and possibly inside the capsid during head assembly. Detection of these proteins is due in part to the exceptional detection limits of the Orbitrap Fusion Lumos mass spectrometer that was used for the analyses of this mutant and *218*(am101).Erwi.

Surprisingly, we detected gp47 in the *47*(am1) particles, however inspection of its sequence coverage revealed that it ended 16 residues prior to residue Q482 whose codon is mutated to a stop codon (whereas in the wild-type virion there is coverage to K561) (**Figure 4**). In addition, the mass spectra identified in this mutant revealed that the proteolytic processing of the truncated form of gp47 was incomplete relative to that which we observed in the wild-type phage (**Figure 4**). From this we inferred that the *N*-terminal propeptide of gp47 likely has a role in targeting the protein into the prohead, similar to the shorter T4 CTS propeptides. Presumably, the incomplete processing of the gp47 propeptide and/or the absence of its *C*-terminus alters the structure and function(s) of gp47 and produces the non-viable phenotype. Based on our analyses showing that the SPN3US virion effectively assembles as a wild-type particle in *47*(am1), we infer that like the vRNAP subunits, gp47 is an ejection protein. Consistent with such a role, the mature fragment of gp47 has an *N*-terminal transmembrane domain, as predicted by TMHMM

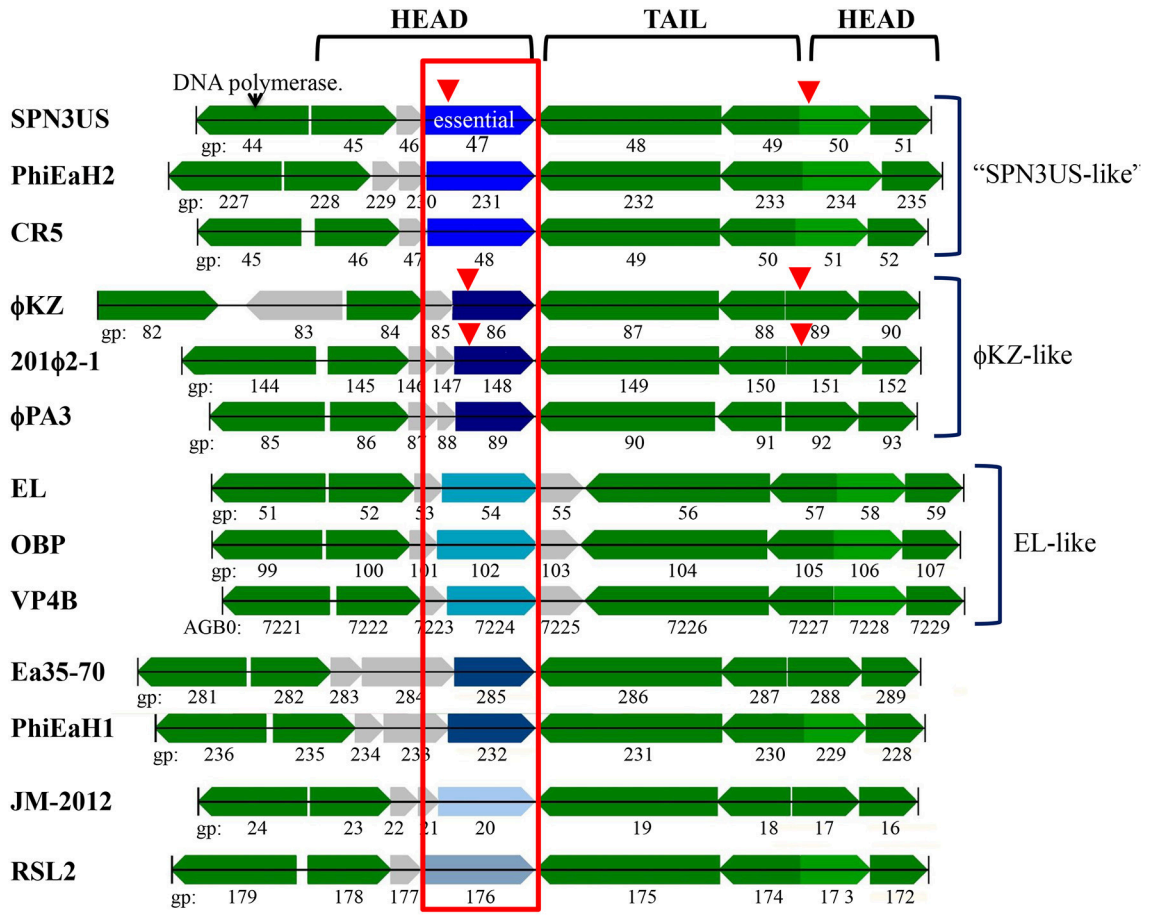


FIGURE 7 | The SPN3US gene region flanking the gene encoding the essential head ejection protein gp47 and corresponding gene regions in related giant phages. Genes in the same gene location as SPN3US_0047 in related giant phages are enclosed in a red box and those shaded the same shade blue were determined to have sequence similarity by Psi-Blast. Other genes with homologs in related phages as determined by Psi-Blast searches are shaded green. SPN3US head or tail genes are indicated. Red arrowhead indicates processed by the prohead protease.

(Krogh et al., 2001) (**Figure 4**) which we speculate interacts with the host membrane during infection. It is interesting to note that if gp47 has a direct host-based function, possibly interacting with the host cell, this may explain the highly diverged set of genes in the equivalent position in related phages which have all evolved based on selective pressures that are a consequence of their own host interactions, akin to the gene plasticity that is seen in tail fiber proteins which have direct interactions with host cell wall components. Our analyses indicate gp47 is an excellent candidate for an inner head protein that is ejected into the host cell, possibly with a role in host takeover.

In contrast to gp47, gp243 is a medium abundance head protein that is not processed by the prohead protease and its general location in the head (i.e., shell vs. internal protein) was unknown. When *243*(am114) was propagated under non-permissive conditions, non-viable particles again were produced; however, a phenotype that is remarkably different to that of the ejection protein mutants described above was observed. The gp243- particles did not survive the first ultracentrifugation in a CsCl step gradient (i.e., no band was produced despite similar yields of sample being loaded onto the gradient), unlike the ejection protein mutants whose particles were stable through both the step and overnight buoyant density gradient ultracentrifugations. Examination by TEM of gp243-particles that had been concentrated by differential centrifugation revealed that there were no intact virions (heads joined to tails) but there were numerous free tails and apparently non-stable head structures (**Figure 8A**). The presence of a structure related to the head is supported by the detection of a band at the appropriate position for the MCP by SDS-PAGE in this sample (**Figure 8B**). Notably, in this sample there were no SDS-PAGE bands observed in the normal positions of the high abundance head proteins gps 53 and 54. From this we infer that gp243 has a function related to the incorporation of gp53 and 54 into the prohead and that without some, or all of gps 243, 53, and 54, the SPN3US head becomes highly unstable. We tentatively assign gp243 as an ejection protein, as it must interact in some manner with gps53 and 54, which are internal head proteins based on their homologs

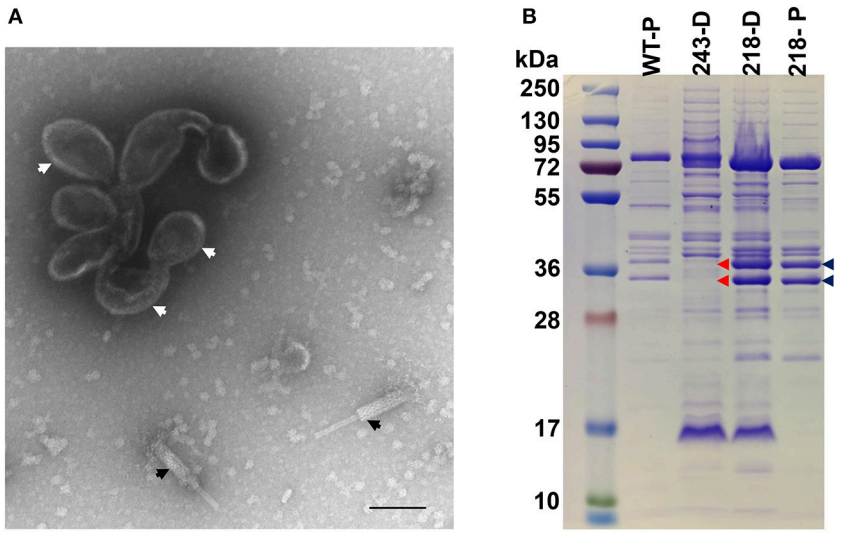


FIGURE 8 | SDS-PAGE profile and morphology of SPN3US mutant *243*(am114). **(A)** Transmission electron microscopy of *243*(am114) after propagation on non-permissive *S. enterica* serovar Typhimurium (strain TT9079) and concentration by differential centrifugation. White arrows indicate disformed, DNA-empty capsid structure, black arrows indicate contracted tails, **(B)** SDS-PAGE gel showing profiles of wild-type (WT) SPN3US and amber mutant phages *243*(am214) and *218*(am101) indicated by their mutated gene names, 243 and 218, respectively, after propagation on the non-permissive strain of *S. enterica* serovar Typhimurium (TT9079). D, indicates particles concentrated by differential centrifugation; P, indicates particles purified by step and buoyant density CsCl ultracentrifugation; Blue arrows indicate the high abundance ejection proteins gp53 and gp54; Red arrows highlight the absence of the high abundance ejection proteins gp53 and gp54 in *243*(am114).

in ϕKZ being components of its large inner body (Thomas et al., 2012).

Identification of a Diverged Homolog to the SPN3US Prohead Protease

A PSI-BLAST search of the SPN3US protease gp245 against the nr and env_nr databases identified the prohead protease homologs in other giant phages, including ϕKZ gp175, the type peptidase for MEROPs family S80 (Rawlings et al., 2012). What was unexpected was that in the second iteration of this search there was a weak match to another SPN3US protein, gp117 (1e-13 in round 4). Notably, the T4 protease gp21 was also identified in this iteration, scoring 2e-16, as were gp21 homologs in many T4-related phages including IME08 (2e-16) and *Synechococcus* phage S-PM2 (1e-34), suggesting the likelihood of non-homologous matches in the profile to be low. The identification of these matches demonstrates how sequence-to-profile based searches have gained power due to increased numbers of sequences in the databases since the ϕKZ protease had to be identified using Hidden Markov Model HMM-based strategies (Thomas et al., 2012).

A reverse PSI-BLAST search from SPN3US gp117 initially found homologs to proteins in phages that infect *E. amylovora,* such as Stratton gp135 and Kwan gp144, none of which were their assigned prohead protease (e.g., Stratton gp267 and Kwan gp271). On examination, these proteins were also identified with comparable scores to gp117 in the PSI-BLAST searches from SPN3US gp245. In the third and later iterations, matches to the biochemically validated prohead proteases of SPN3US and ϕKZ and their homologs in related phages (e.g., 201ϕ2-1 gp268, PhiPA3 gp205, *Vibrio* phage JM-2012 TSMG0080) were drawn into the profile. Further inspection revealed that several phages, in addition to SPN3US and the *Erwinia* phages, had a diverged match to their prohead proteases, such as 201ϕ2-1 gp206 (145 aa), PhiPA3 gp143 (136 aa).

The presence of a paralogous gene to its prohead protease in a phage genome was unprecedented in the literature to our knowledge, so we sought to further test the match between SPN3US gp117 and known prohead proteases using HMM-based strategies. First, an alignment was made between gp117 and 12 homologs from 9 *Erwinia* phages with the Sequence and Alignment Modeling software (SAM) (Hughey and Krogh, 1996; Hughey et al., 2003). (Note that there are two homologs to gp117 in the phages Machina, Caitlin and ChrisDB.) An HMM based on this alignment was scored against a library of all SPN3US proteins, and the prohead protease gp245 had an E-value of 2.4e-09. Conversely, an HMM based on the SPN3US protease gp245 aligned with its homologous proteases in 24 related phages scored gp117 at 1.50e-09. Although the E-values in our searches are inflated as the result from searches of small libraries, the identification of a false positive seems less likely when two different models are each able to find the protein of interest in reverse searches.

To further interrogate the validity of SPN3US gp117 having similarity to gp245 and other known proteases, we used HHpred for its sensitive profile-to-profile (HMM-HMM) based searches (Söding, 2005; Söding et al., 2005). An HHsearch of the gp245 HHM against the T4 gp21 HHM used originally to identify the ϕKZ protease gp175 (Thomas et al., 2012) gave an E-value of 1.1e-11 and an alignment with the three residues of the catalytic

triad of T4 gp21 was found (gp245 residues H-77, S-153, D-178). An HHsearch of the gp117 HHM against this T4 HHM gave an E-value 4e-07 and aligned two of the three catalytic residues in the T4 enzyme to gp117 residues S-126 and D-144. Notably, an HHsearch of the two SPN3US HHMs against one another gave an E-value of 3e-17 and also aligned the catalytic serine and aspartate of gp245 with gp117 residues S-126 and D-144. While there was no direct alignment between gp117 and the catalytic histidine of the two known protease HHMs, there is a histidine seven residues upstream in gp117.

DISCUSSION

The SPN3US Head Structure Is Architecturally T4 gone Rococo

Our analyses of the SPN3US head highlight major themes of head structure and assembly that are conserved between SPN3US, related giant phages, and T4. Critically, the proteins identified in T4 as essential for the formation of its large myoviral capsid shell (major capsid protein and portal) clearly have homologs in SPN3US despite high divergence at the sequence level. SPN3US and related phages all also have homologs to the two essential enzymes in T4 required for head maturation and DNA packaging—the prohead protease and the large terminase protein, respectively (**Figure 9**). Despite the absence of identifiable homologs in T4 for the other SPN3US head protein, the SPN3US head does have numerous shared features with the T4 head, notably internal head proteins that are ejected into the host cell as well as paralogous proteins. The existence of paralogs in both T4 and SPN3US is intriguing as most dsDNA phage genomes do not have paralogs (Kristensen et al., 2011). That both phage genomes do contain paralogous genes is likely the consequence of a shared ancestral replication/recombination pathway as evidenced by diverged homologs in SPN3US to the T4 DNA polymerase (SPN3US gps 18 and 44 which represent a split subunit DNA polymerase as initially identified in phage OBP, Cornelissen et al., 2012) and UvsX (SPN3US gp216).

The detection of proteolysis in eight SPN3US head proteins revealed shared characteristics with the processing of head proteins that occurs in T4. Most notably, in both phages cleavage always occurs after a glutamate residue in a short motif. Our studies also revealed that processing can occur on the *N*-termini and/ or *C*-termini of head proteins in both SPN3US and T4. That both SPN3US and ϕKZ proteases undergo *C*-terminal autoprocessing highlights a need to resolve the relevance of this event in prohead assembly and maturation and/or enzyme activation. Additional biochemical studies of the T4 protease are also needed to elucidate the autocleavage mechanism of gp21 and the role of gp21 in head maturation.

A number of processed SPN3US head proteins showed evidence of multiple processing sites in their propeptide regions. Multiple processing with a substrate protein by the prohead protease is a well-known feature of the T4 MCP propeptide as well as core proteins, such as gp22. We also observed multiple processing sites in ϕKZ head proteins (Thomas et al., 2012; Thomas and Black, 2013). Presumably, this aids in ensuring that these peptides are cleared from the capsid either before or during head expansion/DNA packaging. To comprehensively define the heterogeneity of processing sites for each protein species is beyond current proteomic capabilities. Other analyses are needed to more accurately determine the number and location of the cleavage sites, such as conducted for ϕKZ gp93 (Thomas and Black, 2013). Importantly, the identification of similarities in processing between SPN3US and T4 and also SPN3US and ϕKZ and 201ϕ2-1 provide support for our conclusion that processing by a prohead protease is a conserved, possibly ancient, essential step in head maturation in all related giant phages.

Just as significantly, our analyses of SPN3US have highlighted several major differences that have evolved since T4 and giant phages shared an ancient ancestor. Superficially, these include major variations in head size and structure ($T = 27$ for ϕKZ and SPN3US, vs. $T_{ends} = 13$ for the caps and $T_{mid} = 20$ the prolate capsid of T4, Fokine et al., 2004) in addition to numerous different head proteins in the giant phages (∼50 proteins vs. 13 head proteins in T4, Black et al., 1994). The higher number of head proteins in the giant phages could be attributed to a highly complex, possibly more independent life-cycle of the giant phages, as evidenced by their vRNAPs (Ceyssens et al., 2014) and the large number of head ejection proteins, some of which, such as gp47, likely have roles in host takeover. However, it is also feasible that the high number of head proteins in SPN3US may be, to some extent, a consequence of a genome framework that allows rampant gene duplication and recombination events and that many of these proteins are not essential. We anticipate that this question will be resolved through further analysis of our SPN3US mutant collection to identify all essential head genes.

Additionally, our in-depth analyses of proteolytic processing during head maturation has revealed distinct variations between SPN3US and its relatives, vs. T4. The giant phages have portal proteins with a massive propeptide, unprecedented not only in T4 but any other phage taxon. In addition, the giant phages all have multiple inner head proteins that have propeptides that are much longer than the 10–20 residue propeptide CTS of the T4 internal proteins and Alt. In T4, the CTS functions to ensure that each protein is incorporated into the prohead; the propeptide is then removed from the protein via proteolysis and presumably escapes from the head through small pores in the shell during maturation (Mullaney and Black, 1996). The T4 CTS is so effective at its targeting role that it has been used to package numerous proteins of non-phage origin into T4 heads (Mullaney and Black, 1998; Mullaney et al., 2000). Based on our analyses of SPN3US, particularly of the gp47- mutant into which a *C*-terminally truncated protein was packaged into the head, it is likely the giant phage propeptides have similar functions to that of the T4 CTS sequences. The need for longer propeptides and whether any of the SPN3US propeptides have an additional role, such as core/scaffold formation is yet to be determined although the latter is an important consideration for future studies since a counterpart to the T4 scaffold protein gp22 has not been confirmed in any giant phage.

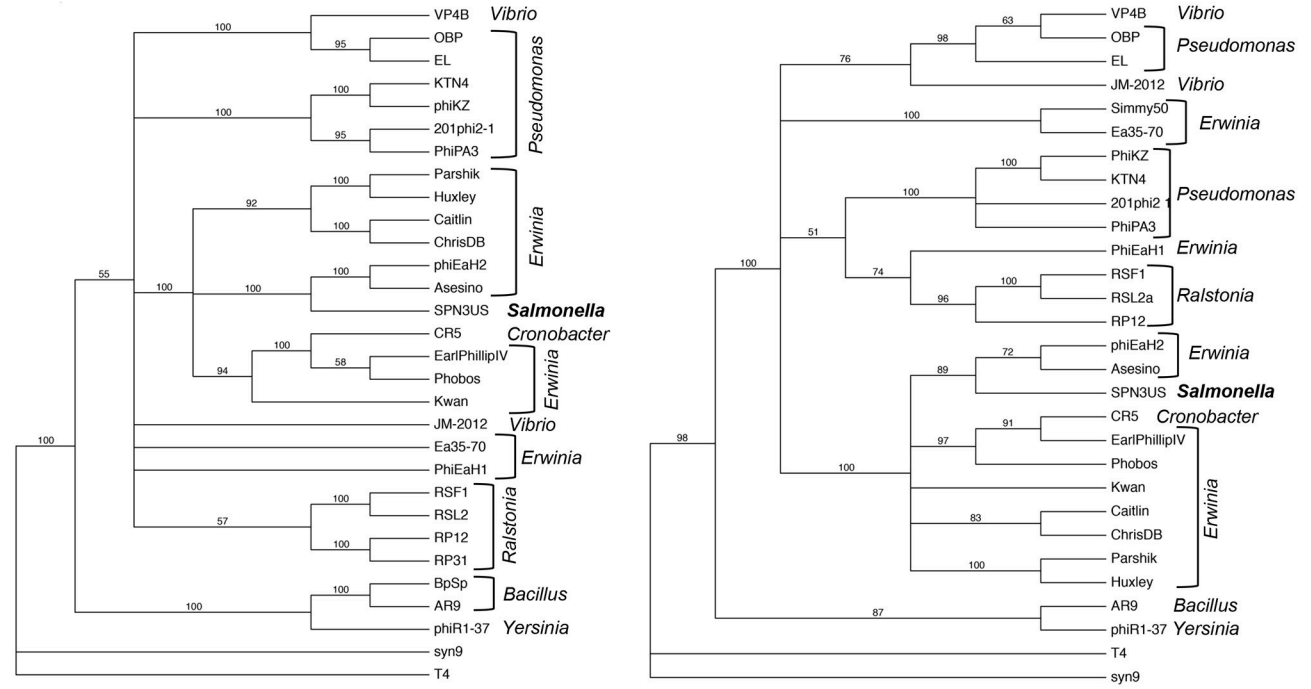


FIGURE 9 | Phylogenetic trees of phage head maturation enzymes in SPN3US and related giant phages. **(A)** Phylogentic tree of the large terminase subunit (TerL), and **(B)** phylogenetic tree of the prohead protease. T4 and the T4-related phage Syn9 proteins were used to root the trees. Phage hosts are indicated.

Giant Phage Head Structure/Function Follows a Virtuosic Evolutionary Pathway

A major goal of our study of SPN3US was to characterize its virion comprehensively as a foundation for further studies that implement it as a genetic model for understanding giant phage biology. As such, this new system has revealed new information regarding the functions and assembly mechanisms of core head proteins, such as the vRNAP for which homologs exist in every related giant phage (Skurnik et al., 2012; Ceyssens et al., 2014; Yakunina et al., 2015). For instance, the SPN3US system has confirmed the essential nature of the vRNAP based on our isolation of mutants in genes encoding three of the subunits, including the recently identified "missing" *C*-terminus of the β′ subunit, and in doing so supported an unprecedented scenario in tailed phage head assembly that this enzyme complex assembles as a multimer prior to incorporation into the prohead. This is truly remarkable when one considers that the packaged vRNAP must then undergo subsequent major conformational rearrangements in the DNA packed capsid to allow for its ejection through the ∼30 Å diameter tail tube into the *Salmonella* cell, where it must then reassemble to be able to transcribe the injected phage DNA. Our vRNAP finding raises the question as to whether the vRNAP multimer is active prior to incorporation into the prohead and highlights a need for further study on this remarkable complex. In addition, our new genetic system has facilitated the characterization of essential head proteins gps 47 and 243, including that the function of gp47 is related to host infection/takeover and that gp243 has a role in the incorporation of members of the paralog family A, for which there are counterparts in all related phages.

Our analyses of SPN3US have revealed unexpected, almost virtuosic, aspects of giant phage head composition and assembly. Regarding composition, the most obvious examples are the head paralog families which show remarkable plasticity in numbers between different phages but also variations in abundance within the same phage. For instance, the numbers of proteins belonging to paralog family A containing the PFAM domain 12699 has been shown to vary between two members (e.g., SPN3US, this study) to seven members (e.g., phiPA3, Cornelissen et al., 2012), while the numbers of paralog family B proteins varies between two (e.g., ϕKZ, Mesyanzhinov et al., 2002) to 20 members in SPN3US. Our estimate of the copy numbers of the processed paralog family A members (>600 copies each per virion) was much higher than our estimates for any of the ϕKZ inner head proteins (Thomas et al., 2012). This was unexpected, as it implies the combined molecular mass of these two 31-kDa proteins in the SPN3US head (>40 MDa) is at least double the estimate of the molecular mass of the ϕKZ IB (Thomas et al., 2012; Wu et al., 2012) and highlights a need for further research to more rigorously quantify internal head protein copy numbers, not only in SPN3US but other related giant phages.

The need for further studies to more accurately quantify the copy numbers of ejection proteins in giant phage heads is additionally underscored by the fact that the SPN3US genome is

∼40 kb shorter than that of ϕKZ. Both SPN3US and ϕKZ likely have the same headful packaging strategy as T4 as their DNAs are packaged to about the same density within their capsids and their large terminase proteins have homology to that of T4, gp17 (**Figure 9**). In T4, DNA packaging concludes when gp17 cleaves the concatemeric DNA by sensing that the capsid is completely full of DNA (Black, 2015). That is, cleavage by the packaging motor is not based on genome length or sequence specificity, but rather reflects the DNA density within the head. Hence, if there were additional, or conversely, less ejection proteins in a phage's capsid, its terminase would accommodate by packaging less or more genomic DNA, respectively. Further research is required to clarify the roles of different ejection proteins in giant phages and also to test our hypothesis that there is a relationship between ejection protein abundance, DNA packaging and genome length in giant phages.

Our studies also highlighted an unexpected plasticity in the proteolysis maturation step among different giant phages. Despite the fact that the SPN3US head is composed of similar numbers of proteins as 201ϕ2-1 and ϕKZ, we found that proteolytic processing occurs during head maturation in only eight SPN3US head proteins, in contrast to the 19 processed proteins in both 201ϕ2-1 and ϕKZ (Thomas et al., 2010, 2012). Initially, we attributed this difference to variations in protein composition between the phages, but based on our correlation of MS sequence coverage and gel migration of individual SPN3US proteins in the wild-type and numbers of mutant proteomes these explanations do not fully account for the reduced number of processed proteins in SPN3US vs. 201ϕ2-1 and ϕKZ.

The variability in the proteolytic processing status of homologous proteins in SPN3US, 201ϕ2-1, and ϕKZ is illustrated in the products of a syntenous head gene region which includes the vRNAP β subunit and protease genes (**Figure 10**). Although the prohead protease in all three phages undergoes auto-proteolysis, the processing status of other proteins from this region is variable. For instance, the 201ϕ2-1 vRNAP β subunit is processed by removal of an *N*-terminal propeptide (cleaved at TFE-275) (Thomas et al., 2010) of similar length to the long propeptide removed from the portal proteins of all three phages (**Figure 10**). Strikingly, our evidence clearly indicates that the SPN3US vRNAP β subunit is not processed (Supplementary Figure 2). We also believe that the ϕKZ homolog gp178 is not processed, but since there were about 10-fold less spectra identified for gp178 than either its 201ϕ2-1 or SPN3US counterparts, this conclusion requires confirmation.

We found no evidence to support proteolytic processing of several other SPN3US proteins although their homologs in other giant phages are processed. For instance, SPN3US essential protein gp243 is not processed (204 total spectra were detected giving an overall coverage of 78%) (Supplementary Figure 3) but 201ϕ2-1 gp271 and ϕKZ gp177 are both cleaved by their prohead proteases at AVE-61 and SVE-60, respectively. Similarly, essential head protein gp214 of SPN3US (Thomas et al., 2016) is not processed (Supplementary Figure 3) whereas both its ϕKZ homolog, gp153, and 201ϕ2-1 homolog, gp238, are processed at SQE-52 and STE-64, respectively.

The variability in the proteolytic processing of head proteins we observed in different giant phages was unexpected because we had assumed that all the internal head proteins of the giant phages would undergo proteolytic processing, because they share essential assembly steps with T4 and all T4 internal head proteins are processed. Also, we had expected that any essential head protein with a conserved function in a giant phage would likely go through the same assembly and maturation processes in related giant phages. That neither expectation held true leads to the conclusion that for numerous giant phage head proteins, there are no negative consequences in terms of protein function/phage viability if the regions considered as propeptides in their counterparts in related phages are not removed by proteolysis. There is probably no better illustration of this than the SPN3US vRNAP β subunit which functions with the long *N*-terminal region still attached, although it is feasible that the retention of this domain affects enzymatic activity/specificity relative to the RNAPs of other giant phages in which it is removed, such as 201ϕ2-1 gp274/3.

A major question arising from our observations of major variations in processing of head proteins in giant phages is "What were the forces that led to these variations?" Did mutations in the protease gene alter enzyme specificity and, therefore, influence which proteins could be processed? Indirect evidence that this may have occurred is that the SPN3US protease has a narrower cleavage sequence specificity (A-X-E) relative to that of its counterparts in 201ϕ2-1 (S/A/G/T-X-E, with 2 A-X-E processing sites) and ϕKZ (S/A/G-X-E, with eight A-X-E processing sites). Additionally, did an event(s) affecting the protease gene alter protease function? Support that such an event, may have occurred can be found in our identification of an extremely diverged match to the SPN3US and T4 proteases, gp117. Notably, the equivalent to the 3′ end of the gp245 protease gene is absent in the gp117 gene, although the downstream gene gp118 is an appropriate length if fused with gp117 to form a protein of almost identical length to that of gp245. In gp245 we infer it is its *C*-terminal region that targets the enzyme into the prohead because it is removed via auto-proteolysis (**Figure 4**). Genetic analyses of T4 showed that if the protease is not incorporated into the prohead, head morphogenesis is effectively frozen and no viable virions are produced (Showe et al., 1976a,b). If we assume that the removal of the *C*-terminal region of a giant phage protease would have a similarly disastrous outcome, then, logically, that phage could only form viable progeny after a protease gene truncation event in one of two ways: 1. if proteolytic processing of proteins was not essential and/or 2. if the phage acquired a version of a protease gene that encoded an active enzyme with similar packaging and substrate specificities as the original enzyme. The latter could occur via a duplication event within the same genome or a recombination event with a related phage. While we can only speculate about the existence of both a protease gene and a potential protease remnant in the SPN3US genome, there is excessive evidence within its genome, and those of related giant phages, that gene splitting, duplication and recombination events have abounded during their evolution making such scenarios

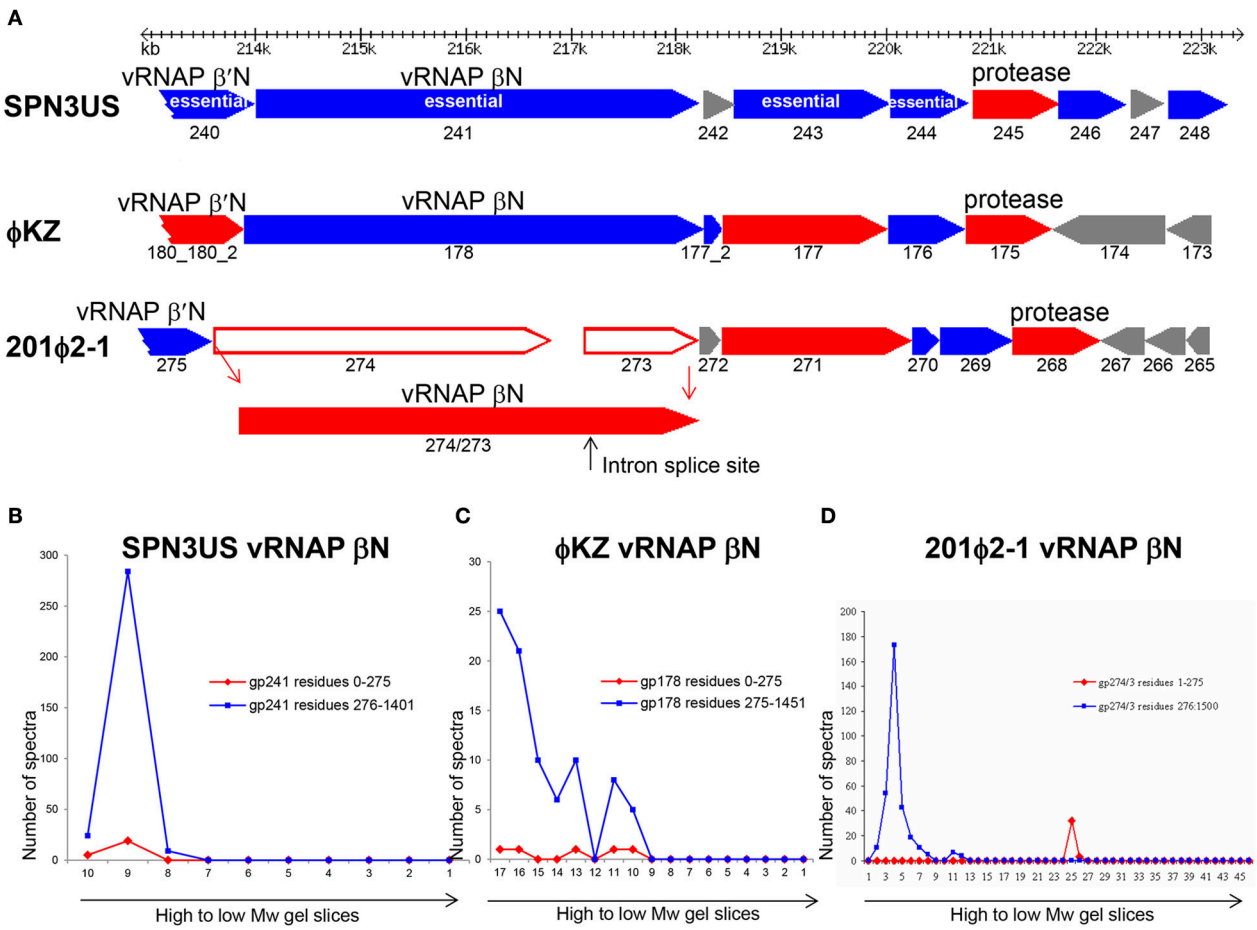


FIGURE 10 | Comparison of proteolytic processing in homologous SPN3US, ϕKZ and 201ϕ2-1 proteins **(A)** Gene region encoding four essential head proteins including the vRNAP βN subunit in SPN3US and corresponding regions in ϕKZ and 201ϕ2-1. Red arrows indicate genes encoding virion proteins processed by the prohead protease, blue arrows indicate genes encoding virion proteins not processed by the prohead protease. Gray arrows indicate non-virion genes. Plot of total mass spectra detected in SDS-PAGE gel slices by mass spectrometry for the RNAP βN subunit of **(B)** SPN3US, **(C)** ϕKZ, and **(D)** 201ϕ2-1.

as described not implausible. In addition, Liu and Mushegian (2004) demonstrated that displacement of protease genes has occurred many times within the order *Caudovirales,* albeit on a broader scale between proteases with different enzymatic specificities, Herpesvirus-like proteases and Clp-like proteases (Liu and Mushegian, 2004).

We conclude that the prohead proteases in both T4-like and giant phages have remarkable functions, cleaving thousands of head proteins in just a few minutes to facilitate a major remodeling of the prohead prior to DNA packaging. Our study highlights that there is still much to be learned about the prohead proteases in both giant phages and T4. However, the variations we have observed in head protein proteolysis between different phages indicate that head maturation has undergone myriad evolutionary events. Consequently, giant phage proteases have likely had a greater impact on giant phage head assembly, structure, composition and possibly even genome length than previously realized.

AUTHOR CONTRIBUTIONS

JT conceived and supervised the project. JT, BA, MD, SM, AB, and LWB performed experiments and data analyses. SW supervised the MS analysis and performed the MS data analysis. LJB and MO assisted with bioinformatics analyses. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Thomas H. Gosnell School of Life Sciences, the College of Science and a GWBC award from the Vice President of Research at RIT (JT), NIH grant AI11676 (LWB) and NIH grant 1S10RR025111-01 (SW).

ACKNOWLEDGMENTS

We acknowledge Phage Biology (BIOL 335) members Mariah Baldwin and Ahmed Tarmizi Abdul Halim for the isolation

of am101 ("SCRAM_3"), Samantha Lomb and Lindsay Smith for the isolation of am84 ("Old Reliable"). In addition we acknowledge Martine Bosch and Adriana Coll De Peña for the isolation of am114. We thank Dr. Stephen C. Hardies for invaluable discussions and allowing use of his bioinformatics resources and Dr. David Lawlor for his reading of the manuscript and helpful comments. We also thank the following individuals: Qin Dan for technical support; Dr. Ru-Ching Hsia for TEM analyses; Kevin Hakala, Sammy Pardo, and Dana Molleur for mass spectrometry analyses; Michelle Zanache, Dr. John Ashton and Jason Myers (University of Rochester Genomics Research Center) for sequencing of mutants and helpful advice; and Dr. Borries Demeler and the UTHSCSA Bioinformatics Center for assistance with computational aspects of the project. Mass spectrometry analyses were conducted at the UTHSCSA Institutional Mass Spectrometry Laboratory. Transmission electron microscopy was performed at the UMB Electron Microscopy Core Imaging Facility. This manuscript is dedicated to the memory of Robert Howlett Thomas.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2017.02251/full#supplementary-material

REFERENCES

Abremski, K., and Black, B. W. (1979). The function of bacteriophage T4 internal protein I in a restrictive strain of *Escherichia coli*. *Virology* 97, 439–447. doi: 10.1016/0042-6822(79)90353-2

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Bair, C. L., Rifat, D., and Black, L. W. (2007). Exclusion of glucosyl-hydroxymethylcytosine DNA containing bacteriophages is overcome by the injected protein inhibitor IPI*. *J. Mol. Biol.* 366, 779–789. doi: 10.1016/j.jmb.2006.11.049

Baker, M. L., Jiang, W., Rixon, F. J., and Chiu, W. (2005). Common ancestry of herpesviruses and tailed DNA bacteriophages. *J. Virol.* 79, 14967–14970. doi: 10.1128/JVI.79.23.14967-14970.2005

Bhunchoth, A., Blanc-Mathieu, R., Mihara, T., Nishimura, Y., Askora, A., Phironrit, N., et al. (2016). Two asian jumbo phages, phiRSL2 and phiRSF1, infect *Ralstonia solanacearum* and show common features of phiKZ-related phages. *Virology* 494, 56–66. doi: 10.1016/j.virol.2016.03.028

Black, L. W. (2015). Old, new, and widely true: the bacteriophage T4 DNA packaging mechanism. *Virology* 479–480, 650–656. doi: 10.1016/j.virol.2015.01.015

Black, L. W., and Abremski, K. (1974). Restriction of phage T4 internal protein I mutants by a strain of *Escherichia coli*. *Virology* 60, 180–191. doi: 10.1016/0042-6822(74)90375-4

Black, L. W., and Rao, V. B. (2012). Structure, assembly, and DNA packaging of the bacteriophage T4 head. *Adv. Virus Res.* 82, 119–153. doi: 10.1016/B978-0-12-394621-8.00018-2

Black, L. W., Showe, M. K., and Steven, A. C. (1994). "Morphogenesis of the T4 head," in *Molecular Biology of Bacteriophage T4,* ed J. D. Karam (Washington, DC: ASM Press), 218–258.

Black, L. W., and Thomas, J. A. (2012). "Condensed genome structure," in *Viral Molecular Machines. Advances in Experimental Medicine and Biology*, eds M. G. Rossmann and V. B. Rao (New York, NY: Springer), 469–487.

Ceyssens, P.-J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., et al. (2014). Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. *J. Virol.* 88, 10501–10510. doi: 10.1128/JVI.01347-14

Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K., et al. (2017). Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355, 194–197. doi: 10.1126/science.aal2130

Cheng, H., Shen, N., Pei, J., and Grishin, N. V. (2004). Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. *Protein Sci.* 13, 2260–2269. doi: 10.1110/ps.04726004

Cornelissen, A., Hardies, S. C., Shaburova, O. V., Krylov, V. N., Mattheus, W., Kropinski, A. M., et al. (2012). Complete genome sequence of the giant virus OBP and comparitive genome analysis of the diverse PhiKZ-related phages. *J. Virol.* 86, 1844–1852. doi: 10.1128/JVI.06330-11

Danis-Wlodarczyk, K., Vandenheuvel, D., Jang, H. B., Briers, Y., Olszak, T., Arabski, M., et al. (2016). A proposed integrated approach for the preclinical evaluation of phage therapy in Pseudomonas infections. *Sci. Rep.* 6:28115. doi: 10.1038/srep28115

De Smet, J., Zimmermann, M., Kogadeeva, M., Ceyssens, P.-J., Vermaelen, W., Blasdel, B., et al. (2016). High coverage metabolomics analysis reveals phage-specific alterations to *Pseudomonas aeruginosa* physiology during infection. *ISME J.* 10**,** 1823–1835. doi: 10.1038/ismej.2016.3

Donelli, G. (1968). Isolamento di un batteriofago di eccezionali dimensioni attivo su B. megatherium. *Atti. Accad. Naz. Lincei Rend. Clas. Sci. Fis. Mat. Nat.* 44, 95–97.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Epstein, R. H., Bolle, A., and Steinberg, C. M. (2012). Amber mutants of bacteriophage T4D: their isolation and genetic characterization. *Genetics* 190, 831–840. doi: 10.1534/genetics.112.138438

Epstein, R. H., Bolle, A., Steinberg, C., Kellenberger, E., Boy de la Tour, E., Chevalley, R., et al. (1963). Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harb. Symp. Quant. Biol.* 28, 375–394. doi: 10.1101/SQB.1963.028.01.053

Fokine, A., Battisti, A., Bowman, V., Efimov, A., Kurochkina, L., Chipman, P., et al. (2007). CryoEM Study of the Pseudomonas bacteriophage phiKZ. *Structure* 15, 1099–1104. doi: 10.1016/j.str.2007.07.008

Fokine, A., Battisti, A. J., Kostyuchenko, V. A., Black, L. W., and Rossmann, M., G (2006). Cryo-EM structure of a bacteriophage T4 gp24 bypass mutant: the evolution of pentameric vertex proteins in icosahedral viruses. *J. Struct. Biol.* 154, 255–259. doi: 10.1016/j.jsb.2006.01.008

Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B., and Rossmann, M. G. (2004). Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6003–6008. doi: 10.1073/pnas.0400444101

Fokine, A., Kostyuchenko, V. A., Efimov, A. V., Kurochkina, L. P., Sykilinda, N. N., Robben, J., et al. (2005a). A three-dimensional cryo-electron microscopy structure of the bacteriophage phiKZ head. *J. Mol. Biol.* 352, 117–124. doi: 10.1016/j.jmb.2005.07.018

Fokine, A., Leiman, P. G., Shneider, M. M., Ahvazi, B., Boeshans, K. M., Steven, A. C., et al. (2005b). Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci.U.S.A.* 102, 7163–7168. doi: 10.1073/pnas.0502164102

Fokine, A., and Rossmann, M. G. (2016). Common evolutionary origin of procapsid proteases, phage tail tubes, and tubes of bacterial type vi secretion systems. *Structure* 24, 1928–1935. doi: 10.1016/j.str.2016.08.013

Gleghorn, M. L., Davydova, E. K., Rothman-Denes, L. B., and Murakami, K. S. (2008). Structural basis for DNA-hairpin promoter recognition by the bacteriophage N4 virion RNA polymerase. *Mol. Cell* 32, 707–717. doi: 10.1016/j.molcel.2008.11.010

Hardies, S. C., Thomas, J. A., Black, L., Weintraub, S. T., Hwang, C. Y., and Cho, B. C. (2016). Identification of structural and morphogenesis genes of

Pseudoalteromonas phage phiRIO-1 and placement within the evolutionary history of Podoviridae. *Virology* 489, 116–127. doi: 10.1016/j.virol.2015.12.005

Harper, D. R., and Morales, S. (2012). Bacteriophage therapy: practicability and clinical need meet in the multidrug-resistance era. *Future Microbiol.* 7, 797–799. doi: 10.2217/fmb.12.58

Hendrix, R. W. (2009). "Jumbo bacteriophages," in *Lesser Known Large dsDNA Viruses. Current Topics in Microbiology and Immunology*, ed J. Van Etten (Berlin; Heidelberg: Springer), 229–240.

Hughey, R., Karplus, K., and Krogh, A. (2003). *SAM: Sequence Alignment and Modeling Software System*. Technical Report UCSC-CRL-99-11, University of California, Santa Cruz, CA. Available online at: https://compbio.soe.ucsc.edu/papers/sam_doc/sam_doc.html

Hughey, R., and Krogh, A. (1996). Hidden Markov models for sequence analysis: extension and analysis of the basic method. *Comput. Appl. Biosci.* 12, 95–107. doi: 10.1093/bioinformatics/12.2.95

Jang, H. B., Fagutao, F. F., Nho, S. W., Park, S. B., Cha, I. S., Yu, J. E., et al. (2013). Phylogenomic network and comparative genomics reveal a diverged member of the ΦKZ-related group, marine Vibrio phage ΦJM-2012. *J. Virol.* 87, 12866–12878. doi: 10.1128/JVI.02656-13

Karplus, K., Barrett, C., and Hughey, R. (1998). Hidden markov models for detecting remote protein homologies. *Bioinformatics* 14, 846–856. doi: 10.1093/bioinformatics/14.10.846

Kazmierczak, K. M., Davydova, E. K., Mustaev, A. A., and Rothman-Denes, L. B. (2002). The phage N4 virion RNA polymerase catalytic domain is related to single-subunit RNA polymerases. *EMBO J.* 21, 5815–5823. doi: 10.1093/emboj/cdf584

Keller, B., and Bickle, T. A. (1986). The nucleotide sequence of gene 21 of bacteriophage T4 coding for the prohead protease. *Gene* 49, 245–251. doi: 10.1016/0378-1119(86)90285-4

Kraemer James, A., Erb Marcella, L., Waddling Christopher, A., Montabana Elizabeth, A., Zehr Elena, A., Wang, H., et al. (2012). A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral dna within the host cell. *Cell* 149, 1488–1499. doi: 10.1016/j.cell.2012.04.034

Kristensen, D. M., Cai, X., and Mushegian, A. (2011). Evolutionarily conserved orthologous families in phages are relatively rare in their prokaryotic hosts. *J. Bacteriol.* 193, 1806–1814. doi: 10.1128/JB.01311-10

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Krylov, V. N., Dela Cruz, D. M., Hertveldt, K., and Ackermann, H. W. (2007). "PhiKZ-like viruses", a proposed new genus of myovirus bacteriophages. *Arch. Virol.* 152, 1955–1959. doi: 10.1007/s00705-007-1037-7

Krylov, V. N., Smirnova, T. A., Minenkova, I. B., Plotnilova, T. G., Zhazikov, I. Z., and Khrenova, E. A. (1984). Pseudomonas bacteriophage phiKZ contains an inner body in its capsid. *Can. J. Microbiol.* 30, 758–762. doi: 10.1139/m84-116

Krylov, V., Shaburova, O., Krylov, S., and Pleteneva, E. (2012). A genetic approach to the development of new therapeutic phages to fight *pseudomonas aeruginosa* in wound infections. *Viruses* 5, 15–53. doi: 10.3390/v5010015

Lecoutere, E., Ceyssens, P.-J., Miroshnikov, K. A., Mesyanzhinov, V. V., Krylov, V. N., Noben, J.P., et al. (2009). Identification and comparative analysis of the structural proteomes of PhiKZ and EL, two giant *Pseudomonas aeruginosa* bacteriophages. *Proteomics* 9, 3215–3219. doi: 10.1002/pmic.200800727

Lee, J.-H., Bai, J., Shin, H., Kim, Y., Park, B., Heu, S., et al. (2016). A novel bacteriophage targeting *Cronobacter sakazakii* is a potential biocontrol agent in foods. *Appl. Environ. Microbiol.* 82, 192–201. doi: 10.1128/AEM.01827-15

Lee, J.-H., Shin, H., Kim, H., and Ryu, S. (2011). Complete genome sequence of Salmonella bacteriophage SPN3US. *J. Virol.* 85, 13470–13471. doi: 10.1128/JVI.06344-11

Leskinen, K., Blasdel, B., Lavigne, R., and Skurnik, M. (2016). RNA-Sequencing reveals the progression of phage-host interactions between ϕR1-37 and *Yersinia enterocolitica*. *Viruses* 8:111. doi: 10.3390/v8040111

Lhuillier, S., Gallopin, M., Gilquin, B., Brasilès, S., Lancelot, N., Letellier, G., et al. (2009). Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8507–8512. doi: 10.1073/pnas.0812407106

Liu, J., and Mushegian, A. (2004). Displacements of prohead protease genes in the late operons of double-stranded-DNA bacteriophages. *J. Bacteriol.* 186, 4369–4375. doi: 10.1128/JB.186.13.4369-4375.2004

Meczker, K., Domotor, D., Vass, J., Rakhely, G., Schneider, G., and Kovacs, T. (2014). The genome of the *Erwinia amylovora* phage PhiEaH1 reveals greater diversity and broadens the applicability of phages for the treatment of fire blight. *FEMS Microbiol. Lett.* 350, 25–27. doi: 10.1111/1574-6968.12319

Mesyanzhinov, V. V., Robben, J., Grymonprez, B., Kostyuchenko, V. A., Bourkaltseva, M. V., Sykilinda, N. N., et al. (2002). The genome of bacteriophage phiKZ of *Pseudomonas aeruginosa*. *J. Mol. Biol.* 317, 1–19. doi: 10.1006/jmbi.2001.5396

Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Ruger, W. (2003). Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.* 67, 86–156. doi: 10.1128/MMBR.67.1.86-156.2003

Movva, N. R., Semon, D., Meyer, C., Kawashima, E., Wingfield, P., Miller, J. L., et al. (1990). Cloning and nucleotide sequence of the *Salmonella typhimurium* pepM gene. *Mol. Gen. Genet.* 223, 345–348.

Mullaney, J. M., and Black, L. W. (1996). Capsid targeting sequence targets foreign proteins into bacteriophage T4 and permits proteolytic processing. *J. Mol. Biol.* 261, 372–385. doi: 10.1006/jmbi.1996.0470

Mullaney, J. M., and Black, L. W. (1998). Activity of foreign proteins targeted within the bacteriophage T4 head and prohead: implications for packaged DNA structure. *J. Mol. Biol.* 283, 913–929. doi: 10.1006/jmbi.1998.2126

Mullaney, J. M., Thompson, R. B., Gryczynski, Z., and Black, L., W (2000). Green fluorescent protein as a probe of rotational mobility within bacteriophage T4. *J. Virol. Methods* 88, 35–40. doi: 10.1016/S0166-0934(00)00166-X

Orlova, E. V., Gowen, B., Dröge, A., Stiege, A., Weise, F., Lurz, R., et al. (2003). Structure of a viral DNA gatekeeper at 10 A resolution by cryo-electron microscopy. *EMBO J.* 22, 1255–1262. doi: 10.1093/emboj/cdg123

Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P., and Yu, H., D (2008). PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in *Pseudomonas* and *Burkholderia* spp. and other bacteria. *Appl. Environ. Microbiol.* 74, 7422–7426. doi: 10.1128/AEM.01369-08

Rawlings, N. D., Barrett, A. J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 40, D343–D350. doi: 10.1093/nar/gkr987

Rifat, D., Wright, N. T., Varney, K. M., Weber, D. J., and Black, L. W. (2008). Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target. *J. Mol. Biol.* 375, 720–734. doi: 10.1016/j.jmb.2007.10.064

Rixon, F. J., and Schmid, M., F (2014). Structural similarities in DNA packaging and delivery apparatuses in Herpesvirus and dsDNA bacteriophages. *Curr. Opin. Virol.* 5, 105–110. doi: 10.1016/j.coviro.2014.02.003

Rossmann, M. G., Arisaka, F., Battisti, A. J., Bowman, V. D., Chipman, P. R., Fokine, A., et al. (2007). From structure of the complex to understanding of the biology. *Acta Crystallogr. D. Biol. Crystallogr.* 63 (Pt 1), 9–16. doi: 10.1107/S0907444906047330

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Serwer, P., Hayes, S. J., Thomas, J. A., and Hardies, S. C. (2007). Propagating the missing bacteriophages: a large bacteriophage in a new class. *Virol. J.* 4:21. doi: 10.1186/1743-422X-4-21

Serwer, P., Hayes, S. J., Thomas, J., Demeler, B., and Hardies, S. C. (2009). "Isolation of novel large and aggregating bacteriophages," in *Bacteriophages: Methods and Protocols. Methods in Molecular Biology*, eds M. Clokie and A. M. Kropinski (New York, NY: Humana Press), 55–66.

Serwer, P., Hayes, S. J., Zaman, S., Lieman, K., Rolando, M., and Hardies, S. C. (2004). Improved isolation of undersampled bacteriophages: finding of distant terminase genes. *Virology* 329, 412–424. doi: 10.1016/j.virol.2004.08.021

Showe, M. K., Isobe, E., and Onorato, L. (1976a). Bacteriophage T4 prehead proteinase: I. Purification and properties of a bacteriophage enzyme which cleaves the capsid precursor proteins. *J. Mol. Biol.* 107, 35–54. doi: 10.1016/S0022-2836(76)80016-2

Showe, M. K., Isobe, E., and Onorato, L. (1976b). Bacteriophage T4 prehead proteinase: II. Its cleavage from the product of gene 21 and regulation in phage-infected cells. *J. Mol. Biol.* 107, 55–69. doi: 10.1016/S0022-2836(76)80017-4

Skurnik, M., Hyytiäinen, H. J., Happonen, L. J., Kiljunen, S., Datta, N., Mattinen, L., et al. (2012). Characterization of the genome, proteome, and structure of yersiniophage ϕR1-37. *J. Virol.* 86, 12625–12642. doi: 10.1128/JVI.01783-12

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21, 951–960. doi: 10.1093/bioinformatics/bti125

Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248. doi: 10.1093/nar/gki408

Sokolova, O. S., Shaburova, O. V., Pechnikova, E. V., Shaytan, A. K., Krylov, S. V., Kiselev, N. A., et al. (2014). Genome packaging in EL and Lin68, two giant phiKZ-like bacteriophages of *P. aeruginosa*. *Virology* 468–470, 472–478. doi: 10.1016/j.virol.2014.09.002

Sousa, R., and Mukherjee, S. (2003). T7 RNA polymerase. *Prog. Nucleic Acid Res. Mol. Biol.* 73, 1–41. doi: 10.1016/S0079-6603(03)01001-8

Sun, L., Zhang, X., Gao, S., Rao, P. A., Padilla-Sanchez, V., Chen, Z., et al. (2015). Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nat. Commun.* 6:7548. doi: 10.1038/ncomms8548

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analyses Using Parsimony (*and Other Methods). Version 4*. Sunderland, MA: Sinauer Associates.

Sycheva, L. V., Shneider, M. M., Sykilinda, N. N., Ivanova, M. A., Miroshnikov, K. A., and Leiman, P. G. (2012). Crystal structure and location of gp131 in the bacteriophage phiKZ virion. *Virology* 434, 257–264. doi: 10.1016/j.virol.2012.09.001

Thomas, J. A., Benítez Quintana, A. D., Bosch, M. A., Coll De Pe-a, A., Aguilera, E., Coulibaly, A., et al. (2016). Identification of essential genes in the Salmonella phage SPN3US reveals novel insights into giant phage head structure and assembly. *J. Virol.* 90, 10284–10298. doi: 10.1128/JVI.01492-16

Thomas, J. A., and Black, L. W. (2013). Mutational analysis of the *Pseudomonas aeruginosa* myovirus ϕKZ morphogenetic protease gp175. *J. Virol.* 87, 8713–8725. doi: 10.1128/JVI.01008-13

Thomas, J. A., Rolando, M. R., Carroll, C. A., Shen, P. S., Belnap, D. M., Weintraub, S. T., et al. (2008). Characterization of *Pseudomonas chlororaphis* myovirus 201ϕ2-1 via genomic sequencing, mass spectrometry, and electron microscopy. *Virology* 376, 330–338. doi: 10.1016/j.virol.2008.04.004

Thomas, J. A., S. C. H., Rolando, M., Hayes, S. J., and Lieman, K., Carroll, C. A., et al. (2007). Complete genomic sequence and mass spectrometric analysis of highly diverse, atypical *Bacillus* thuringiensis phage 0305ϕ8-36. *Virology* 368, 405–421. doi: 10.1016/j.virol.2007.06.043

Thomas, J. A., Weintraub, S. T., Hakala, K., Serwer, P., and Hardies, S. C. (2010). Proteome of the large Pseudomonas myovirus 201ϕ2-1: delineation of proteolytically processed virion proteins. *Mol. Cell. Proteomics* 9, 940–951. doi: 10.1074/mcp.M900488-MCP200

Thomas, J. A., Weintraub, S. T., Wu, W., Winkler, D. C., Cheng, N., Steven, A. C., et al. (2012). Extensive proteolysis of head and inner body proteins by a morphogenetic protease in the giant *Pseudomonas aeruginosa* phage ϕKZ. *Mol. Microbiol.* 84, 324–339. doi: 10.1111/j.1365-2958.2012.08025.x

Van den Bossche, A., Hardwick, S. W., Ceyssens, P.-J., Hendrix, H., Voet, M., Dendooven, T., et al. (2016). Structural elucidation of a novel mechanism for the bacteriophage-based inhibition of the RNA degradosome. *Elife* 5:e16413. doi: 10.7554/eLife.16413

Wingfield, P., Graber, P., Turcatti, G., Movva, N. R., Pelletier, M., Craig, S., et al. (1989). Purification and characterization of a methionine-specific aminopeptidase from Salmonella typhimurium. *Eur. J. Biochem.* 180, 23–32. doi: 10.1111/j.1432-1033.1989.tb14610.x

Wu, W., Thomas, J. A., Cheng, N., Black, L. W., and Steven, A. C. (2012). Bubblegrams reveal the inner body structure of ϕKZ. *Science* 335, 182. doi: 10.1126/science.1214120

Yagubi, A. I., Castle, A. J., Kropinski, A. M., Banks, T. W., and Svircev, A. (2014). Complete genome sequence of *Erwinia amylovora* bacteriophage vB_EamM_Ea35-70. *Genome Announc.* 2:e00413-14. doi: 10.1128/genomeA.00413-14

Yakunina, M., Artamonova, T., Borukhov, S., Makarova, K. S., Severinov, K., and Minakhin, L. (2015). A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage. *Nucleic Acids Res.* 43, 10411–10420. doi: 10.1093/nar/gkv1095

Zehr Elena, A., Kraemer James, A., Erb Marcella, L., Coker Joanna, K. C., Montabana Elizabeth, A., Pogliano, J., et al. (2014). The structure and assembly mechanism of a novel three-stranded tubulin filament that centers phage DNA. *Structure* 22, 539–548. doi: 10.1016/j.str.2014.02.006

Zybailov, B., Coleman, M. K., Florens, L., and Washburn, M. P. (2005). Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal. Chem.* 77, 6218–6224. doi: 10.1021/ac050846r

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 November 2018
doi: 10.3389/fmicb.2018.02798

Ubiquitin–Proteasome System Is Required for Efficient Replication of Singapore Grouper Iridovirus

Xiaohong Huang[1], Shina Wei[1], Songwei Ni[2], Youhua Huang[1] and Qiwei Qin[1,3]**

[1] College of Marine Sciences, South China Agricultural University, Guangzhou, China, [2] Key Laboratory of Tropical Marine Bio-Resources and Ecology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China, [3] Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

The ubiquitin–proteasome system (UPS) serves as the major intracellular pathway for protein degradation and plays crucial roles in several cellular processes. However, little is known about the potential actions of the UPS during fish virus infection. In this study, we elucidated the possible roles of UPS in the life cycle of Singapore grouper iridovirus (SGIV); a large DNA virus that usually causes serious systemic diseases with high mortality in groupers. Data from transcriptomic analysis of differentially expressed genes illustrated that expression of 65 genes within the UPS pathway, including ubiquitin encoding, ubiquitination, deubiquitination, and proteasome, were up- or down-regulated during SGIV infection. Using different proteasome inhibitors, inhibition of the proteasome decreased SGIV replication *in vitro*, accompanied by inhibition of virus assembly site formation, and viral gene transcription and protein transportation. Over-expression of ubiquitin partly rescued the inhibitory effect of ubiquitin inhibitor on SGIV replication, suggesting that UPS was required for fish iridovirus infection *in vitro*. Viral or host proteins regulated by proteasome inhibition during SGIV infection were investigated with two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Sixty-two differentially expressed proteins, including 15 viral and 47 host proteins, were identified after SGIV infection. The host proteins were involved in ubiquitin-mediated protein degradation, metabolism, cytoskeleton, macromolecular biosynthesis, and signal transduction. Among them, 11 proteins were negatively regulated upon MG132 treatment during SGIV infection. This is believed to be the first study to provide evidence that UPS was essential for fish virus infection and replication.

Keywords: iridovirus, ubiquitin, proteasome, viral replication, grouper

OPEN ACCESS

Edited by:
Jonatas Abrahao, Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Ludmila Karen Dos Santos Silva, Délégation Alsace (CNRS), France
Eric Roberto Guimarães Rocha Aguiar, Federal University of Bahia, Brazil
Alice Abreu Torres, University of Cambridge, United Kingdom

****Correspondence:***
Youhua Huang
huangyh@scau.edu.cn
Qiwei Qin
qinqw@scau.edu.cn

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

Received: *05 August 2018*
Accepted: *31 October 2018*
Published: *26 November 2018*

Citation:
Huang X, Wei S, Ni S, Huang Y and Qin Q (2018) Ubiquitin–Proteasome System Is Required for Efficient Replication of Singapore Grouper Iridovirus. Front. Microbiol. 9:2798. doi: 10.3389/fmicb.2018.02798

INTRODUCTION

The ubiquitin-proteasome system (UPS) is the major intracellular protein degradation pathway and plays crucial roles in a variety of fundamental cellular processes, including regulation of gene transcription, cell cycle progression, autophagy, development and differentiation, and modulation of the immune and inflammatory responses (Glickman and Ciechanover, 2002; Wertz and Dixit, 2008; Zhao and Goldberg, 2016). There is increasing evidence that the UPS is required for viral

infection by affecting viral entry, gene transcription, assembly, release and immune evasion (Banks et al., 2003; Wang et al., 2016; Casorla-Pérez et al., 2017). To the best of our knowledge, DNA viruses, as well as RNA viruses from different hosts, including mammals, insects, and plants, exploit the UPS system at various stages of the viral life cycle (Camborde et al., 2010; Contin et al., 2011; Katsuma et al., 2011; Greene et al., 2012; Wang et al., 2016). The proteasome machinery seems to play opposing roles during viral infection. On the one hand, proteasome inhibition with bortezomib leads to increased susceptibility to lymphocytic choriomeningitis virus or coronavirus infection *in vivo* (Basler et al., 2009; Raaben et al., 2010). On the other hand, inhibition of proteasome activity prevents viral DNA replication and the formation of virus assembly sites during vaccinia virus (VACV) replication (Satheshkumar et al., 2009). Inhibition of proteasome activity also reduces Kaposi's sarcoma-associated herpesvirus (KSHV) entry into endothelial cells and intracellular trafficking (Greene et al., 2012). Therefore, exploration of the molecular mechanism by which the UPS regulates viral replication will provide an alternative potential target for antiviral therapy.

Singapore grouper iridovirus (SGIV), a novel member of the genus *Ranavirus*, family *Iridoviridae*, was first isolated from diseased groupers. SGIV infection causes >90% mortality in groupers and sea bass (Qin et al., 2001). Our previous studies demonstrated that SGIV infection in grouper cells induces non-apoptotic cell death, and mitogen-activated protein kinase (MAPK) signaling pathways, including extracellular signal-regulated kinase, p38 MAPK, and c-Jun N-terminal kinase signaling, which are involved in viral replication (Huang et al., 2011a,b). Genome annotation of SGIV reveals that some potential viral gene products, including ubiquitin (ORF102L) and predicted E3 ubiquitin ligase (ORF146L), might be involved in the regulation of the UPS during SGIV infection (Song et al., 2004). Transcriptome analysis of SGIV-infected grouper spleen shows that several genes associated with ubiquitin-mediated proteolysis are up- or down-regulated in response to SGIV infection, suggesting that the UPS plays important roles in SGIV infection (Huang Y.H. et al., 2011). However, the molecular mechanism underlying the regulatory effects of UPS on SGIV replication remain uncertain.

In this study, we explored the importance of the UPS in SGIV infection using different proteasome inhibitors. Moreover, viral or cellular proteins regulated by the UPS pathway during SGIV infection were investigated with two-dimensional gel electrophoresis (2-DE) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). This study is believed to be the first to show molecular evidence that the UPS is involved in fish iridovirus infection, providing new clues to understanding the fish-virus interaction.

MATERIALS AND METHODS

Materials

Proteasome inhibitors MG132 (carbobenzoxy-L-leucyl-L-leucyl-L-leucinal) and lactacystin were purchased from Sigma. Bortezomib and ubiquitin-activating enzyme (E1) inhibitor (Pyr-41) were purchased from Selleckchem. These inhibitors were dissolved in dimethyl sulfoxide (DMSO), and their cytotoxicity on grouper spleen (GS) cells was determined using trypan blue exclusion dye staining.

Cells and Virus

The GS cell line used in this study was established in our laboratory (Huang et al., 2009). GS cells were cultured in Leibovitz's L-15 supplemented with 10% fetal bovine serum (Gibco) and kept in incubators at 25°C. SGIV was propagated in GS cells and stored at −80°C. For the inhibition experiments, GS cells were pretreated with DMSO or various concentrations of inhibitors for 2 h, and then infected with SGIV at multiplicity of infection (MOI) of two for indicated times.

To explore which steps of SGIV replication are affected by proteasome disruption, virus production was determined after treatment with MG132 at different time points during SGIV infection as described previously (Chen et al., 2008). MG132 treatment was carried out at different time points during infection, and then cells were washed to remove MG132 for further incubation until 24 h. In detail, GS cells were pre-treated with MG132 for 2 h, and replaced with normal medium in "P" group. In "TH" or "DMSO" group, MG132 or DMSO was present in the medium throughout virus infection. In addition, MG132 was present in the medium only for 0–6 h, 6–12 h, 12–18 h, and 18–24 h in "0–6 h," "6–12 h," "12–18 h," and "18–24 h" groups, respectively. Finally, the whole cell lysate at indicated time points were collected and determined for virus titration as described below.

Viral Titer Assay

Viral titers were determined on monolayers of GS cells in 96-well plates with 50% tissue culture infective dose ($TCID_{50}$) assay as described previously (Huang et al., 2011a). SGIV was serially diluted 10-fold and overlaid on ∼95% confluent monolayers of GS cells in 96-well plates and incubated for 1 h. After removing the medium containing virus, cells were washed with fresh medium three times. Finally, cells were incubated with fresh medium and cultured at 25°C for 5 days. The cytopathic effect (CPE) was observed under microscopy and virus titer ($TCID_{50}$/ml) was calculated according to Reed and Muench (1938). The results were expressed as means of three independent experiments. The statistical significances were determined with Student's *t*-test. The significance level was defined as $p < 0.05$ (*).

RNA Sequencing and Analysis

To explores the expression profiles of host genes in response to SGIV infection, RNA sequencing was carried out in SGIV infected GS cells. In brief, mock- or SGIV-infected GS cells (12, 24, and 48 h p.i.) in triplicate flasks were collected and total RNA was extracted using the mirVana miRNA Isolation Kit (Ambion) following the manufacturer's protocol. The libraries were constructed using TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, United States) and then sequenced on the Illumina sequencing platform (HiSeqTM 2500). After the initial assembly, the differentially expressed genes (DEGs)

related to UPS was analyzed as described previously (Gao and Chen, 2018). The changes of target genes were analyzed using the expression levels in SGIV-infected cells compared to those in mock-infected cells at indicated time points.

Electron Microscopy

DMSO- or MG132-treated GS cells were infected with SGIV and harvested at 24 h post-infection (h p.i.). Sample preparation was performed as previously described (Huang et al., 2011a). After washing with phosphate-buffered saline (PBS), cells were post-fixed in 1% osmium tetroxide for 1 h, and then dehydrated in graded ethanol. The cells were embedded in EPON resin. Sections were double stained with uranyl acetate and lead citrate. The ultrathin sections were examined in a JEM-1400 electron microscopy (Jeol) at 120 kV.

Immunofluorescence Assays

GS cells were grown on coverslips in six-well plates, and then infected with SGIV at MOI 2 in the presence or absence of 10 μM MG132. At 24 h p.i., the infected cells were fixed in 4% paraformaldehyde and permeabilized with absolute alcohol for 7 min at −20°C. After washing with PBS, the cells were blocked with 2% bovine serum albumin for 30 min, and then incubated with primary antibodies (anti-VP19 serum 1:100) for 2 h at room temperature. The coverslips were washed with PBS, followed by incubation with the secondary antibody, fluorescein-isothiocyanate-conjugated goat anti-mouse IgG (1:100, Pierce) for 1 h. Nuclei were stained with 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI). Samples were observed under an inverted fluorescence microscope (Leica).

Quantitative Real-Time PCR Analysis (qRT-PCR)

To confirm the effects of MG132 on viral gene expression, qRT-PCR was used to evaluate the relative RNA expression of several genes. The specific PCR primers for the viral genes are described previously (Huang et al., 2011b). Total cell RNA was extracted from DMSO- or MG132-treated infected cells at 6, 12, and 24 h p.i. RNA extraction was performed using SV Total RNA Isolation Kit (Promega). RNA reverse transcription was carried out using ReverTra Ace qPCR RT Kit (Toyobo) according to the manufacture's instruction. qRT-PCR was performed in a Roche 480 Real Time Detection System (Roche, Germany) using the SYBR Green Real-time PCR Kit (Toyobo) as described previously (Huang et al., 2011b). Each assay was performed in triplicate, and β-actin was chosen as the internal control. The data are representative of three independent experiments. The statistical significances were determined with Student's *t*-test. The significance level was defined as $p < 0.05$ (*).

Western Blot Analysis

At 6, 12, and 24 h p.i., DMSO- or MG132-treated infected cells were harvested, and the pellets were resuspended in 1× lysis buffer (Cell Signaling Technology). SDS–PAGE and western blotting were performed as described previously (Huang et al., 2013b). Equal amounts of protein were subjected to SDS–PAGE and then transferred to polyvinylidene difluoride membranes. After blocking with 5% non-fat dry milk, the membranes were incubated with the primary antibodies for 2 h at room temperature, including anti-VP86 (1:1000), anti-VP136 (1:1000), anti-VP72 (1:1500), and anti-VP19 serum (1:1500). After washing with Tris buffer, the membranes were incubated for 1 h with the HRP-Goat anti-Mouse IgG (1:1000). Simultaneously, internal controls were performed by detecting β-actin protein. Immunoreactive bands were visualized with diaminobenzidine.

2-DE Analysis and Protein Identification

To analyze the differentially expressed proteins regulated by MG132 treatment during SGIV infection, 2-DE was performed as described previously (Zhao et al., 2010). GS cells were pretreated with either DMSO or 10 μM MG132 for 2 h, and then infected with SGIV at MOI 2. Cells were harvested and lysed in lysis buffer [20 mM Tris, 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 2 mM TBP, 0.2% IEF buffer], then protein concentrations were determined using a modified Bradford assay. All samples were stored at −80°C prior to electrophoresis. 2-DE was carried out using Immobiline strips (pI range, 3–10; GE Healthcare, Piscataway, NJ, United States), with proteins being separated according to charge, and subsequently molecular weight. The gels were stained with Coomassie brilliant blue G-250. Differentially expressed protein spots were excised from the gels for MS analysis. MALDI-TOF spectra were calibrated using trypsin autolysis peptide signals and matrix ion signals as described by Liu et al. (2013). Protein identification using peptide mass fingerprinting was performed using the MASCOT search tool ([1]Matrix Science Ltd., London, United Kingdom). Protein identification with a score >85 was regarded as positive identification.

RESULTS

UPS-Related Genes Were Differently Expressed During SGIV Infection

To unravel crucial cellular factors involved in SGIV replication, the differentially expressed cellular genes during virus infection were identified with RNA-Seq analysis. Sixty-four grouper genes related to the UPS components were differently regulated at the different stages of SGIV infection. These genes participated in various aspects of the UPS, including ubiquitination, deubiquitination and proteasome degradation. For example, the expression levels of proteasome subunit α (PSMA)2, PSMA3, ubiquitin-conjugating enzyme (UBE)2A, RING finger protein (RFP)37, ubiquitin, small ubiquitin-related modifier (SUMO)2, and ubiquitin carboxyl-terminal hydrolase (USP)1 were significantly up-regulated during SGIV infection. In contrast, the majority of genes were significantly down-regulated during SGIV infection, including ubiquitin-like modifier-activating enzyme (UBA)1, UBE2N, SUMO1,

[1]http://www.matrixscience.com/

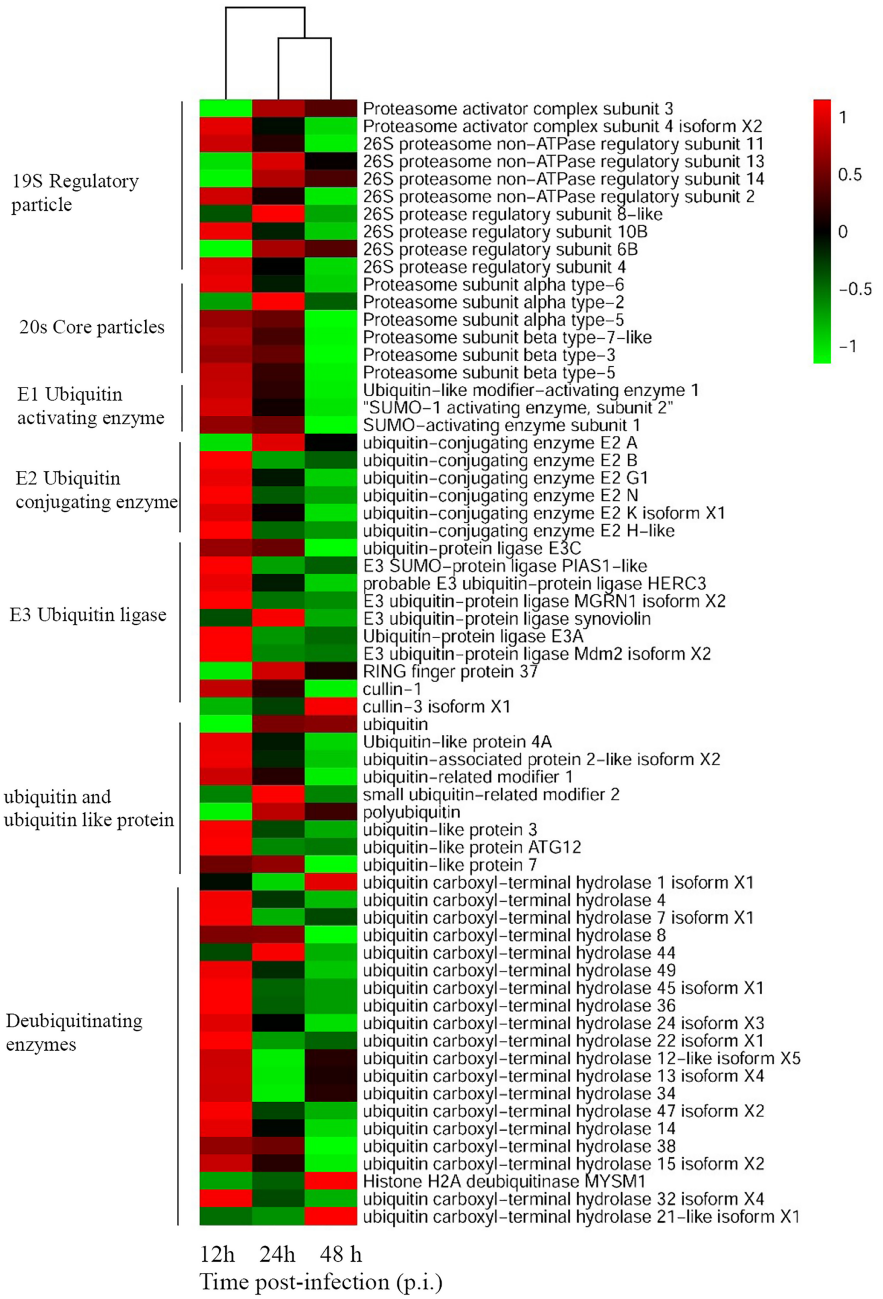


FIGURE 1 | Heatmap representing gene expression of DEGs enriched to UPS during SGIV infection. Red color represents increased expression while green color represents decreased expression. Each column (12, 24, and 48 h) represents the expression levels in SGIV infected GS cells (12, 24, and 48 h) compared to mock-infected GS cells at the corresponding time points.

ubiquitin-like protein (UBL)3, USP36, USP4, and others listed in **Figure 1**.

Multiple Proteasome Inhibitors Decreased SGIV Infection *in vitro*

To determine whether UPS was essential for SGIV infection, three structurally unrelated proteasome inhibitors, MG-132, lactacystin and bortezomib, were used to inhibit proteasome activity during SGIV infection. We evaluated the cytotoxic effects of these inhibitors on GS cells, and selected the optimal concentration of MG132 (10 μM), lactacystin (10 μM), and bortezomib (5 μM) in the following study (**Figure 2A**). After treatment with proteasome inhibitors, a significant delay in the severity of CPE was observed in infected cells treated with MG132, lactacystin or bortezomib, compared with that in DMSO-treated cells (**Figure 2B**). Given that severity of CPE evoked by SGIV is associated with cell viability

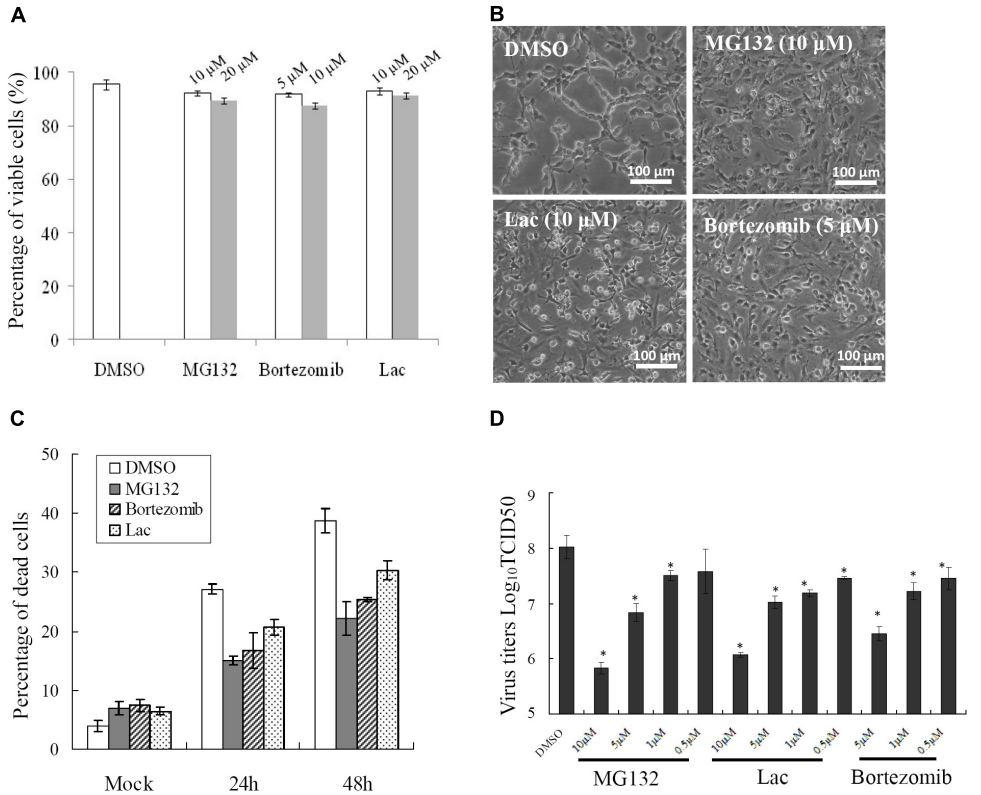


FIGURE 2 | Different proteasome inhibitors reduced viral replication. **(A)** The toxicity of different proteasome inhibitors in GS cells. Effects of proteasome inhibitors on SGIV-induced CPE at 24 h p.i. **(B)**, cell death **(C)**, and viral production **(D)**. $^*p < 0.05$.

(Huang et al., 2011a), we assessed virus-induced cell death under treatment with proteasome inhibitors. Consistently, treatment with these three inhibitors significantly decreased SGIV-induced cell death (**Figure 2C**). The effect of proteasome inhibitors on virus production was also evaluated by viral titer assay. Virus production was significantly reduced at 24 hpi in the presence of 5 or 10 μM MG132, 5 or 10 μM lactacystin or 5 μM bortezomib during infection, suggesting that the effect of proteasome inhibitors on virus production was dose dependent (**Figure 2D**).

Ubiquitin-Activating Enzyme E1 and Ubiquitin Were Involved in SGIV Infection

Ubiquitin-activating enzyme E1 is one of the important components of the UPS, thus, the effect of E1 inhibitor PYR-41 on SGIV infection was also evaluated by viral titer assay. Virus production was significantly decreased in the presence of 10 or 20 μM PYR-41 (non-toxic to GS cells, data not shown) during infection (**Figure 3**). To determine whether inhibition of SGIV replication by MG132 was partially due to depletion of free ubiquitin, grouper ubiquitin was cloned into pCMV-HA vector as described previously (Karpe and Meng, 2012). qRT-PCR analysis indicated that the expression of ubiquitin increased significantly in recombinant plasmid pHA-EcUb overexpressing cells compared to control vector (pCMV-HA) transfected cells (data not shown). Furthermore, over-expression of pHA-EcUb partially countered the inhibitory effects of MG132, including viral production and gene transcription (**Figures 3B,C**). Thus, we propose that ubiquitination was also necessary for the productive infection of SGIV.

Proteasome Inhibitor Inhibited Viral Gene Transcription and Protein Synthesis

To clarify the dynamic alterations of viral replication after proteasome inhibition, viral gene transcription and protein synthesis, including immediately early (VP86), early (VP136) and two late structural (VP72 and VP19) genes were examined in DMSO- or MG132-treated infected cells. At the transcription level, qRT-PCR indicated that the mRNA transcripts of VP86, VP136, VP72, and VP19 were all reduced significantly at different time points in MG132-treated infected cells comparing with the DMSO-treated cells (**Figure 4A**). Consistently, at the protein synthesis level, the protein products of VP72, VP19, VP136, and VP86 were obviously detected from 6 to 24 hpi during infection in DMSO-treated cells. In contrast, VP72 and VP019 were weakly detected at 6 and 24 hpi, while VP86 and VP136 were even undetectable in MG132-treated cells (**Figure 4B**). Our results indicated that viral gene transcription and protein synthesis

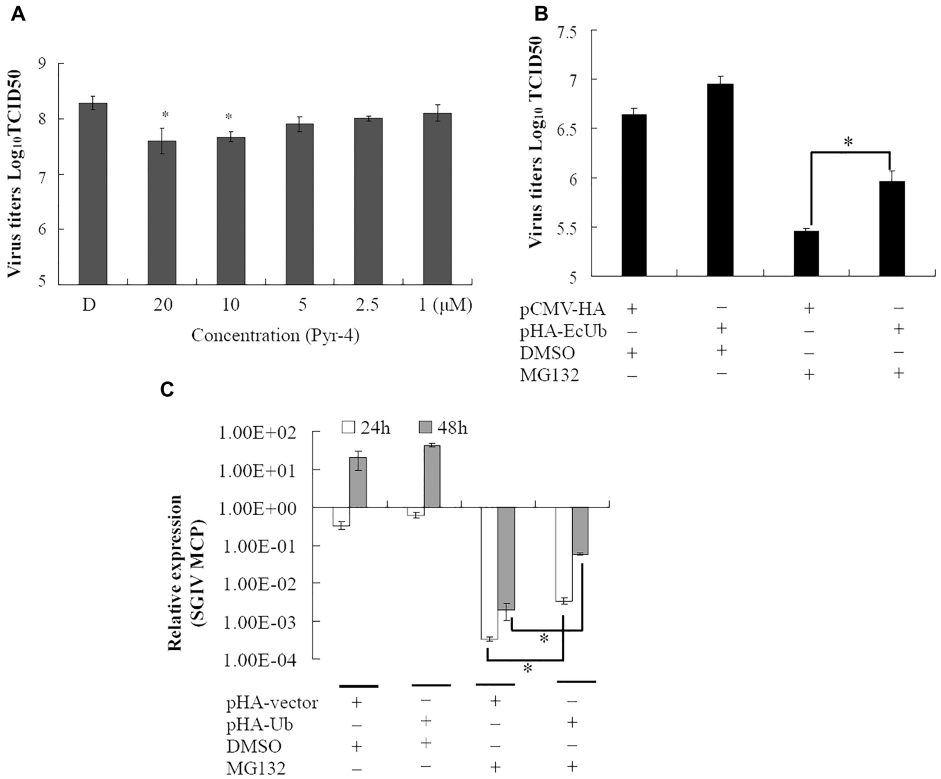


FIGURE 3 | Ubiquitin-activating enzyme (E1) and ubiquitin were involved in virus infection. **(A)** The inhibitory effect of Pyr-41 on viral production was dose dependent. Over-expression of grouper ubiquitin partially countered the effect of MG132 on SGIV production **(B)** and viral gene transcription **(C)**. *$p < 0.05$.

during SGIV infection were severely inhibited by MG132 treatment.

Proteasome Inhibitors Prevented Formation of Viral Factories and Transportation of Viral Proteins

As a large enveloped DNA virus, SGIV replicates and assembles in viral factories that form at pericentriolar sites. Under fluorescence microscopy, viral factories were observed after staining with DAPI during SGIV infection. Many viral assembly sites were observed in DMSO-treated SGIV-infected cells, but few in MG132-treated cells (**Supplementary Figure S1A**). To examine the ultrastructural morphology of viral factories, SGIV-infected DMSO- or MG132-treated cells were prepared for electron microscopy. Numerous viral particles were observed in almost all the cells, and the viral factories were present close to the nucleus in the majority of SGIV-infected DMSO-treated cells at 24 hpi (**Figure 5A**). In contrast, in MG132-treated infected cells, only a few viral particles and no factories were observed.

Cytoplasmic DNA viruses usually concentrate the structural proteins into viral assembly sites at the late stage of infection (Heath et al., 2001; Zhao et al., 2008; Huang et al., 2013a). In this study, SGIV VP19 proteins were found to be mostly localized in the viral factories in DMSO-treated infected cells at 24 h p.i. Green fluorescence spots were randomly distributed in the cytoplasm in the MG132-treated infected cells at 24 h p.i. (**Figure 5B**). Consistently, VP75 proteins also overlapped with viral factories in SGIV-infected DMSO-treated cells at 24 h p.i., and displayed punctate fluorescent spots in MG132-treated infected cells (**Supplementary Figure S1B**). Thus, our results suggested that proteasome inhibition not only prevented transportation of viral proteins, but also affected the formation of viral factories during SGIV infection.

Proteasome Disruption Exerted More Crucial Roles at the Early Stage of SGIV Infection

To delineate the potential mechanisms of proteasome on SGIV infection, reversible inhibitor MG132 was added at different times during SGIV infection as shown in **Figure 6A**. The virus titer assay showed that treatment with MG132 for 0–6 h resulted in significant decrease of virus production (1.6 log-unit reduction compared to DMSO treated cells). The virus titer of the group treated for 6–12 h was 1 log unit lower than that of the control. The groups treated with MG132 from 12 to 18 h p.i., and from 18–24 h p.i. showed a slight decrease of virus titer (**Figure 6B**), suggesting that the addition of

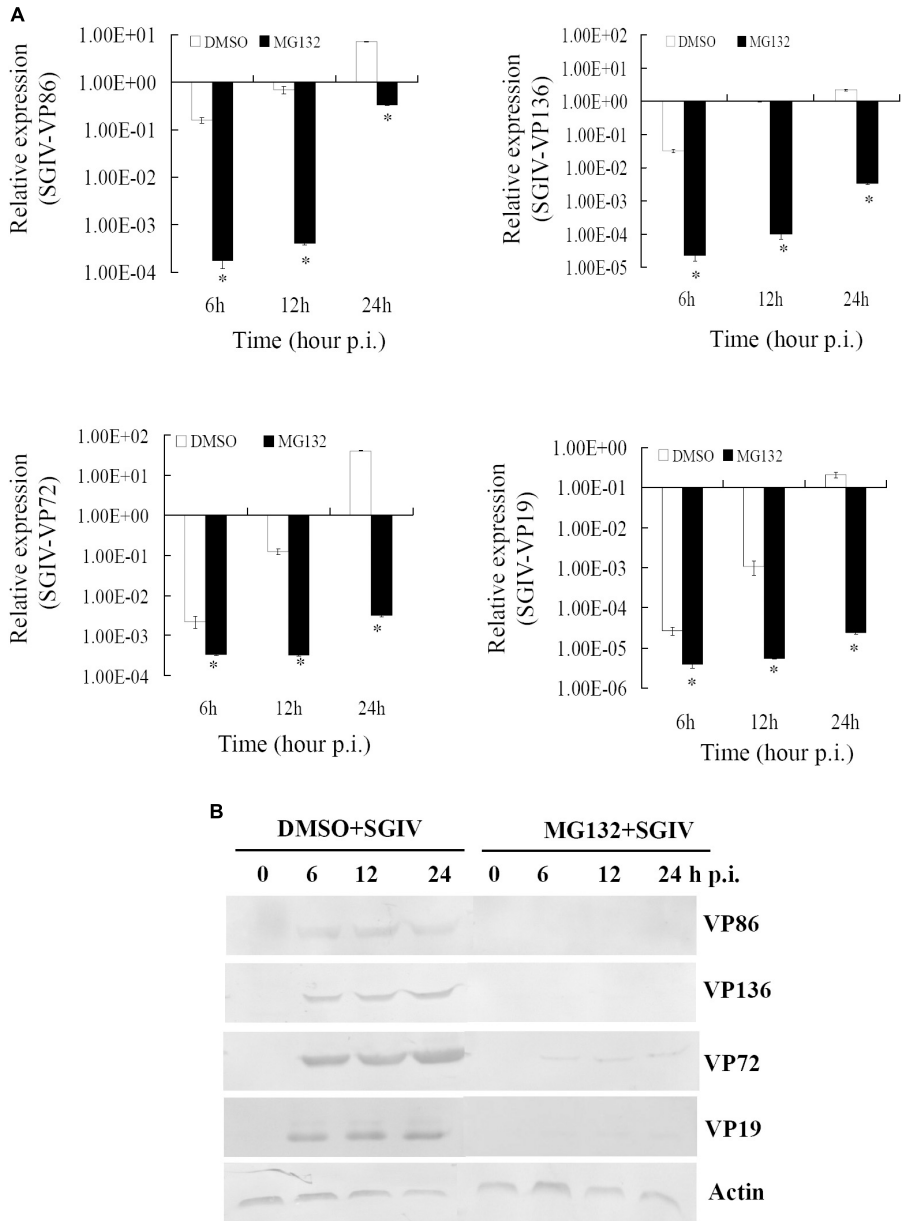


FIGURE 4 | MG132 treatment significantly reduced viral gene transcription and protein synthesis. **(A)** Transcription of viral genes (immediate early gene VP86, early gene VP136, and late genes VP72 and VP19) was decreased by MG132. *$p < 0.05$. **(B)** Synthesis of viral proteins was also impaired by MG132.

MG132 played more important roles at the early stage of SGIV infection.

Proteasome Inhibition Regulated Host Proteins Involved in SGIV Replication

To investigate further the potential mechanism underlying the action of the UPS during SGIV infection, the protein samples collected from SGIV-infected and mock-infected cells in the presence or absence of MG132 were separated using 2-DE. One hundred and thirty protein spots were obviously altered in SGIV-infected cells or MG132-treated SGIV-infected cells. After MS analysis, 62 differentially expressed spots were identified, including 15 viral proteins and 47 host proteins. The identified spots were marked with numbers (**Supplementary Figure S2**), and the retrieved proteins corresponding to each numbered spots are listed in **Supplementary Table S1**. All the identified viral proteins were significantly down-regulated, and only VP67 and VP6 were weakly detectable in MG132-treated infected cells (**Figure 7A**). This implied that viral protein synthesis was severely decreased after proteasome inhibition.

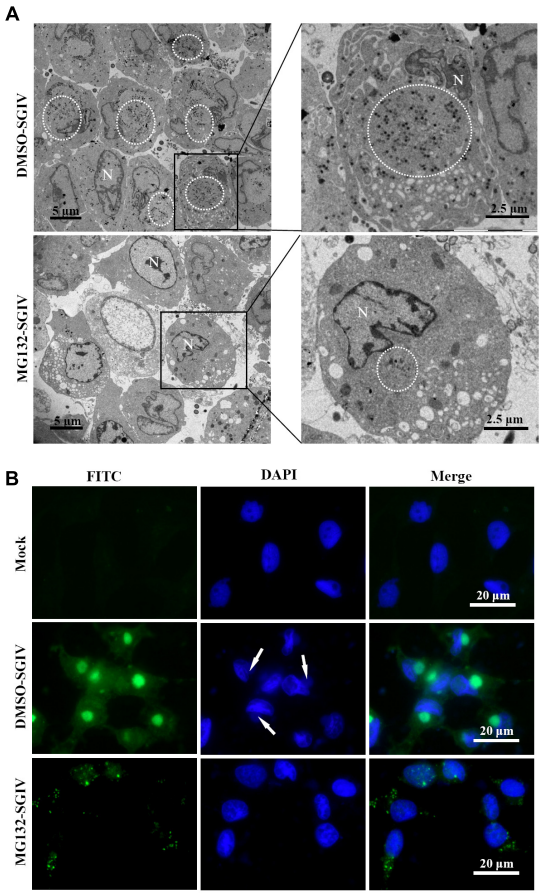


FIGURE 5 | Transport of viral proteins and formation of viral factories were impaired after treatment with MG132. **(A)** Formation of viral factories after MG132 treatment. Circular places show the viral factories. N indicated nucleus. **(B)** Intracellular localization of VP019 after SGIV infection in the absence or presence of MG132. Arrows showed the virus factories.

According to protein functions and subcellular annotations from the Gene Ontology Database, the identified cellular proteins were involved in the cytoskeleton, macromolecular biosynthesis, metabolism, ubiquitin–proteasome pathway, and stress response. Among these regulated proteins, PDLIM1, PFN2, CTSB, and DUT were significantly up-regulated during SGIV infection. Interestingly, these proteins were significantly down-regulated in MG132-treated infected cells compared to mock-infected cells (**Figure 7B**). In contrast, SEPT2, PDB1, UROD, NAS, PDHA1, RpLP0, and SSR4 were significantly down-regulated during SGIV infection, while increased in MG132-treated infected cells compared to mock-infected cells (**Figure 7C**). Thus, our results suggested that certain host proteins involved in SGIV infection were regulated by proteasome inhibition.

DISCUSSION

The UPS can be exploited by different mammalian viruses during their life cycles, including during entry, assembly and release (Harty et al., 2001; Ott et al., 2002; Delboy et al., 2008; Kaspari et al., 2008; Tran et al., 2010; Widjaja et al., 2010; Casorla-Pérez et al., 2017). However, the potential roles of the UPS in fish viral infections remain largely uncertain (Huang et al., 2017). In our study, RNA-Seq based transcriptome analysis of SGIV-infected cells indicated that numerous genes related to the UPS were differentially regulated during SGIV infection. These genes were involved in different aspects of the UPS, including ubiquitination, deubiquitination and proteasome degradation. Proteasome subunit PSMA2, PSMA3, ubiquitin, E3 ubiquitin ligase, RFP37, UBE2A, and deubiquitinating enzyme USP1 were all significantly up-regulated during SGIV infection, suggesting that the UPS was involved in SGIV replication. During dengue virus serotype 2 infection, expression of ubiquitin-activating enzyme E1 (UBE1) and proteasome subunits were increased (Kanlaya et al., 2010). In addition, ubiquitin-conjugating enzyme, 26S proteasome regulatory subunits, and ubiquitin were also differentially regulated by tomato ringspot virus infection (Babu et al., 2008).

Although the UPS plays crucial roles during different viral infections, the underlying mechanisms are different (Delboy et al., 2008; Camborde et al., 2010; Tran et al., 2010; Greene et al., 2012). Proteasome inhibitors block avian reovirus replication at an early stage in the viral life cycle, but do not affect entry and internalization (Chen et al., 2008). The UPS is essential at all stages of human cytomegalovirus infection (Kaspari et al., 2008). In our study, both proteasome inhibitors and ubiquitin-activating enzyme E1 inhibitor delayed CPE progression in SGIV infection and reduced the viral products. The formation of viral factories was also inhibited after proteasome destruction. Vaccinia-virus-infected, MG132-treated cells also lack viral assembly sites within the cytoplasm, which is accompanied by absence of late gene expression (Teale et al., 2009). Over-expression of grouper ubiquitin partly reverses the inhibitory effect of MG132 on SGIV replication, suggesting that ubiquitin also plays crucial roles in SGIV replication, like other mammalian viruses (Si et al., 2008; Karpe and Meng, 2012; Cheng et al., 2014). As two separate arms of the UPS, ubiquitylation, and proteasomal degradation are closely linked and act at different stages (Greene et al., 2012). Therefore, we propose that the UPS is required for fish iridovirus infection *in vitro*.

As a major intracellular protein degradation system, the UPS is involved in a variety of fundamental cellular processes, including regulation of gene transcription and cell signaling, cell cycle, and cell proliferation and differentiation (Yao and Ndoja, 2012). Using 2-DE and MS analysis, we identified 62 differentially expressed proteins, including 15 viral proteins and 47 host proteins after MG132 treatment. Consistent with the data from western blotting, all the identified viral proteins were found in SGIV-infected cells, and almost undetectable in SGIV-infected MG132-treated cells, suggesting that viral protein synthesis were impaired after MG132 treatment. Apart from the viral proteins, certain host cellular proteins involved in different cell processes were regulated by SGIV infection or MG132 treatment. Among them, PDLIM1, PFN2, CTSB, and DUT were significantly up-regulated during SGIV infection, but significantly down-regulated in SGIV-infected MG132-treated cells. In contrast, SEPT2, PDB1,

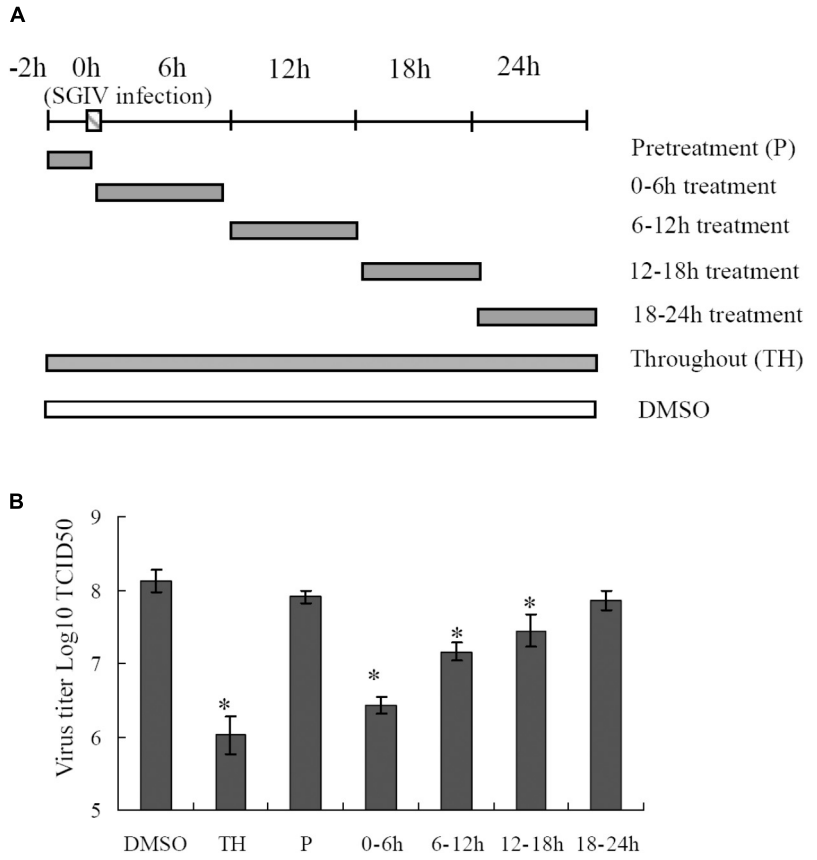


FIGURE 6 | Effects of MG132 on SGIV replication at different stages of virus infection. **(A)** Experimental design for time frame experiments. GS cells were infected with SGIV and treated with MG132 at different times. The cells were washed to remove the drug and further incubated until 24 h. Whole cell lysates were collected and virus production was determined. **(B)** MG132 severely affected the virus replication at an early step during SGIV infection. $^*p < 0.05$.

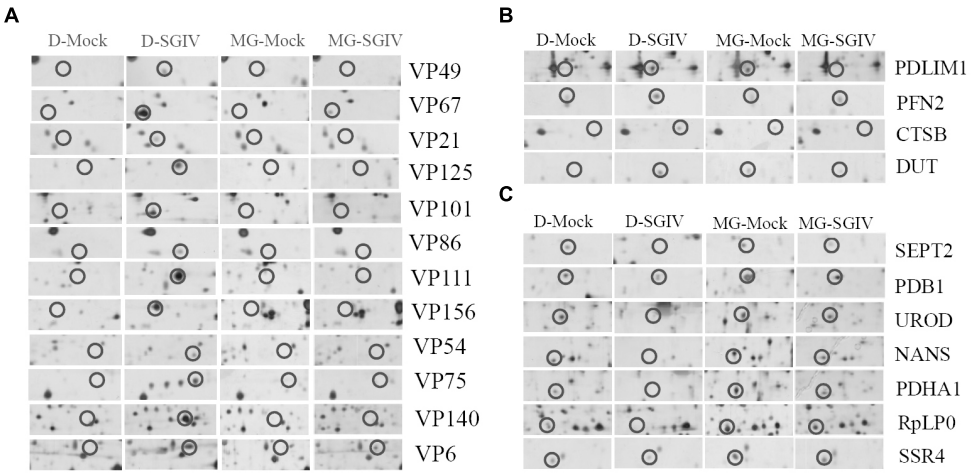


FIGURE 7 | Viral or host proteins were differentially expressed in SGIV-infected cells after treatment with MG132. **(A)** Viral proteins were decreased under MG132 treatment. SGIV infection up-regulated **(B)** or down-regulated **(C)** proteins were impaired by MG132 treatment.

UROD, NAS, PDHA1, RpLP0, and SSR4 were significantly down-regulated during SGIV infection, but only slightly decreased in SGIV-infected MG132-treated cells. It has been reported that CTSB aggravates CVB3-induced viral myocarditis, probably through activating the inflammasome and promoting pyroptosis (Wang et al., 2018). Depletion of SEPT2 in HeLa cells increases

replication of VACV (Beard et al., 2014). Our previous studies also demonstrated that grouper CTSB was involved in SGIV replication and SGIV induced apoptosis (Wei et al., 2014). Whether the action of CTSB on SGIV infection was mediated by UPS still remained uncertain. In addition, PDLIM1 negatively regulates nuclear factor (NF)-κB-mediated signaling, and PFN2 encodes an actin-binding protein involved in endocytosis (Ono et al., 2015; Luscieti et al., 2017). Whether these proteins exerted similar roles in SGIV infection and were regulated by UPS needs further investigation.

In summary, we reported the actions of the UPS in the life cycle of SGIV. Numerous genes related to the UPS were up/down-regulated during SGIV infection, and UPS destruction impaired SGIV replication, as demonstrated by the decrease in viral gene transcription, protein synthesis and formation assembly sites. MG132 treatment regulated certain cellular proteins that were involved in viral infection, suggesting that the UPS plays crucial roles during SGIV infection via regulation of host proteins. Thus, our study provides new insights into understanding the underlying molecular mechanism of the UPS during SGIV infection.

AUTHOR CONTRIBUTIONS

XH and YH carried out the main experiments, analyzed the data, and drafted the manuscript. SW and SN performed the viral titer assay and participated in the qRT-PCR experiments. YH and QQ designed the experiments and reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (31372566 and 31472309), National Key R&D Program of China (2017YFC1404504), and China Agriculture Research System (CARS-47-G16). The mass spectrometry was performed in Instrumental Analysis & Research Center, Sun Yat-sen University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02798/full#supplementary-material

FIGURE S1 | (A) MG132 reduced formation of viral assembly sites. Viral assembly sites in SGIV-infected, DMSO- or MG132-treated GS cells were stained with DAPI and observed under fluorescence microscopy. **(B)** Transport of viral structural protein VP75 was impaired by MG132. Arrows indicate the viral assembly sites.

FIGURE S2 | Protein expression profiles of the SGIV-infected and mock-infected cells in the presence or absence of MG132. The differentially expressed protein spots are marked with numbers for identification.

TABLE S1 | Altered proteins in Mock or SGIV-infected GS cells.

REFERENCES

Babu, M., Griffiths, J. S., Huang, T. S., and Wang, A. (2008). Altered gene expression changes in *Arabidopsis* leaf tissues and protoplasts in response to Plum pox virus infection. *BMC Genomics* 9:325. doi: 10.1186/1471-2164-9-325

Banks, L., Pim, D., and Thomas, M. (2003). Viruses and the 26S proteasome: hacking into destruction. *Trends. Biochem. Sci.* 28, 452–459. doi: 10.1016/S0968-0004(03)00141-5

Basler, M., Lauer, C., Beck, U., and Groettrup, M. (2009). The proteasome inhibitor bortezomib enhances the susceptibility to viral infection. *J. Immunol.* 183, 6145–6150. doi: 10.4049/jimmunol.0901596

Beard, P. M., Griffiths, S. J., Gonzalez, O., Haga, I. R., Pechenick Jowers, T., Reynolds, D. K., et al. (2014). A loss of function analysis of host factors influencing vaccinia virus replication by RNA interference. *PLoS One* 9:e98431. doi: 10.1371/journal.pone.0098431

Camborde, L., Planchais, S., Tournier, V., Jakubiec, A., Drugeon, G., Lacassagne, E., et al. (2010). The ubiquitin-proteasome system regulates the accumulation of Turnip yellow mosaic virus RNA-dependent RNA polymerase during viral infection. *Plant. Cell.* 22, 3142–3152. doi: 10.1105/tpc.109.072090

Casorla-Pérez, L. A., López, T., López, S., and Arias, C. F. (2017). The ubiquitin-proteasome system is necessary for the efficient replication of human astrovirus. *J. Virol.* 92, e1809–e1817. doi: 10.1128/JVI.01809-17

Chen, Y. T., Lin, C. H., Ji, W. T., Li, S. K., and Liu, H. J. (2008). Proteasome inhibition reduces avian reovirus replication and apoptosis induction in cultured cells. *J. Virol. Methods* 151, 95–100. doi: 10.1016/j.jviromet.2008.03.016

Cheng, S., Yan, W., Gu, W., and He, Q. (2014). The ubiquitin-proteasome system is required for the early stages of porcine circovirus type 2 replication. *Virology* 45, 198–204. doi: 10.1016/j.virol.2014.03.028

Contin, R., Arnoldi, F., Mano, M., and Burrone, O. R. (2011). Rotavirus replication requires a functional proteasome for effective assembly of viroplasms. *J. Virol.* 85, 2781–2792. doi: 10.1128/JVI.01631-10

Delboy, M. G., Roller, D. G., and Nicola, A. V. (2008). Cellular proteasome activity facilitates herpes simplex virus entry at a postpenetration step. *J. Virol.* 82, 3381–3390. doi: 10.1128/JVI.02296-07

Gao, E. B., and Chen, G. (2018). Micropterus salmoides rhabdovirus (MSRV) infection induced apoptosis and activated interferon signaling pathway in largemouth bass skin cells. *Fish Shellfish Immunol.* 76, 161–166. doi: 10.1016/j.fsi.2018.03.008

Glickman, M. H., and Ciechanover, A. (2002). The ubiquitin-proteasome protelytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82, 373–428. doi: 10.1152/physrev.00027.2001

Greene, W., Zhang, W., He, M., Witt, C., Ye, F., and Gao, S. J. (2012). The ubiquitin/proteasome system mediates entry and endosomal trafficking of Kaposi's sarcoma-associated herpesvirus in endothelial cells. *PloS Pathog.* 8:e1002703. doi: 10.1371/journal.ppat.1002703

Harty, R. N., Brown, M. E., McGettigan, J. P., Wang, G., Jayakar, H. R., Huibregtse, J. M., et al. (2001). Rhabdoviruses and the cellular ubiquitin-proteasome system: a budding interaction. *J. Virol.* 75, 10623–10629. doi: 10.1128/JVI.75.22.10623-10629.2001

Heath, C. M., Windsor, M., and Wileman, T. (2001). Aggresomes resemble sites specialized for virus assembly. *J. Cell Biol.* 153, 449–455. doi: 10.1083/jcb.153.3.449

Huang, R., Zhang, J., Zhu, G., He, J., and Xie, J. (2017). The core ubiquitin system of mandarin fish, *Siniperca chuatsi*, can be utilized by infectious spleen and kidney necrosis virus. *Fish Shellfish Immunol.* 70, 293–301. doi: 10.1016/j.fsi.2017.09.017

Huang, X., Gong, J., Huang, Y., Ouyang, Z., Wang, S., Chen, X., et al. (2013a). Characterization of an envelope gene VP19 from Singapore grouper iridovirus. *Virol J.* 10:354. doi: 10.1186/1743-422X-10-354

Huang, X., Huang, Y., Cai, J., Wei, S., Ouyang, Z., and Qin, Q. (2013b). Molecular cloning, expression and functional analysis of ISG15 in orange-spotted grouper, *Epinephelus coioides*. *Fish Shellfish Immunol.* 34, 1094–1102. doi: 10.1016/j.fsi.2013.01.010

Huang, X. H., Huang, Y. H., Ouyang, Z. L., Xu, L. X., Yan, Y., Cui, H., et al. (2011a). Singapore grouper iridovirus, a large DNA virus, induces nonapoptotic cell death by a cell type dependent fashion and evokes ERK signaling. *Apoptosis* 16, 831–845. doi: 10.1007/s10495-011-0616-y

Huang, X. H., Huang, Y. H., OuYang, Z. L., Cai, J., Yan, Y., and Qin, Q. W. (2011b). Roles of stress-activated protein kinases in the replication of Singapore grouper iridovirus and regulation of the inflammatory responses in grouper cells. *J. Gen. Virol.* 92, 1292–1301. doi: 10.1099/vir.0.029173-0

Huang, X. H., Huang, Y. H., Sun, J. J., Han, X., and Qin, Q. W. (2009). Characterization of two grouper *Epinephelus akaara* cell lines: application to studies of Singapore grouper iridovirus (SGIV) propagation and virus–host interaction. *Aquaculture* 292, 172–179. doi: 10.1016/j.aquaculture.2009.04.019

Huang, Y. H., Huang, X. H., Yan, Y., Cai, J., Ouyang, Z. L., Cui, H. C., et al. (2011). Transcriptome analysis of orange-spotted grouper (*Epinephelus coioides*) spleen in response to Singapore grouper iridovirus. *BMC Genomics* 12:556. doi: 10.1186/1471-2164-12-556

Kanlaya, R., Pattanakitsakul, S. N., Sinchaikul, S., Chen, S. T., and Thongboonkerd, V. (2010). The ubiquitin-proteasome pathway is important for dengue virus infection in primary human endothelial cells. *J. Proteome. Res.* 9, 4960–4971. doi: 10.1021/pr100219y

Karpe, Y. A., and Meng, X. J. (2012). Hepatitis E virus replication requires an active ubiquitin-proteasome system. *J. Virol.* 86, 5948–5952. doi: 10.1128/JVI.07039-11

Kaspari, M., Tavalai, N., Stamminger, T., Zimmermann, A., Schilf, R., and Bogner, E. (2008). Proteasome inhibitor MG132 blocks viral DNA replication and assembly of human cytomegalovirus. *FEBS Lett.* 582, 666–672. doi: 10.1016/j.febslet.2008.01.040

Katsuma, S., Tsuchida, A., Matsuda-Imai, N., Kang, W., and Shimada, T. (2011). Role of the ubiquitin-proteasome system in Bombyx mori nucleopolyhedrovirus infection. *J. Gen. Virol.* 92(Pt 3), 699–705. doi: 10.1099/vir.0.027573-0

Liu, X. J., Kang, L. Q., Liu, Y. J., Li, H., and Peng, X. (2013). Characterization of the Edwardsiella tarda proteome in response to different environmental stresses. *J. Proteomics.* 80, 320–333. doi: 10.1016/j.jprot.2013.01.022

Luscieti, S., Galy, B., Gutierrez, L., Reinke, M., Couso, J., Shvartsman, M., et al. (2017). The actin-binding protein profilin 2 is a novel regulator of iron homeostasis. *Blood* 130, 1934–1945. doi: 10.1182/blood-2016-11-754382

Ono, R., Kaisho, T., and Tanaka, T. (2015). PDLIM1 inhibits NF-κB-mediated inflammatory signaling by sequestering the p65 subunit of NF-κB in the cytoplasm. *Sci. Rep.* 5:18327. doi: 10.1038/srep18327

Ott, D. E., Coren, L. V., Sowder, R. C. II, Adams, J., Nagashima, K., and Schubert, U. (2002). Equine infectious anemia virus and the ubiquitin-proteasome system. *J. Virol.* 76, 3038–3044. doi: 10.1128/JVI.76.6.3038-3044.2002

Qin, Q. W., Lam, T. J., Sin, Y. M., Shen, H., Chang, S. F., Ngoh, G. H., et al. (2001). Electron microscopic observations of a marine fish iridovirus isolated from brown-spotted grouper, *Epinephelus tauvina*. *J. Virol. Methods.* 98, 17–24. doi: 10.1016/S0166-0934(01)00350-0

Raaben, M., Grinwis, G. C., Rottier, P. J., and de Haan, C. A. (2010). The proteasome inhibitor Velcade enhances rather than reduces disease in mouse hepatitis coronavirus-infected mice. *J. Virol.* 84, 7880–7885. doi: 10.1128/JVI.00486-10

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty percent endpoints. *Am. J. Epidemiol.* 27, 493–497. doi: 10.1093/oxfordjournals.aje.a118408

Satheshkumar, P. S., Anton, L. C., Sanz, P., and Moss, B. (2009). Inhibition of the ubiquitin-proteasome system prevents vaccinia virus DNA replication and expression of intermediate and late genes. *J. Virol.* 83, 2469–2479. doi: 10.1128/JVI.01986-08

Si, X., Gao, G., Wong, J., Wang, Y., Zhang, J., and Luo, H. (2008). Ubiquitination is required for effective replication of coxsackievirus B3. *PLoS One* 3:e2585. doi: 10.1371/journal.pone.0002585

Song, W. J., Qin, Q. W., Qiu, J., Huang, C. H., Wang, F., and Hew, C. L. (2004). Functional genomics analysis of Singapore grouper iridovirus: complete sequence determination and proteomic analysis. *J. Virol.* 78, 12576–12590. doi: 10.1128/JVI.78.22.12576-12590.2004

Teale, A., Campbell, S., Van Buuren, N., Magee, W. C., Watmough, K., Couturier, B., et al. (2009). Orthopoxviruses require a functional ubiquitin-proteasome system for productive replication. *J. Virol.* 83, 2099–2108. doi: 10.1128/JVI.01753-08

Tran, K., Mahr, J. A., and Spector, D. H. (2010). Proteasome subunits relocalize during human cytomegalovirus infection, and proteasome activity is necessary for efficient viral gene transcription. *J. Virol.* 84, 3079–3093. doi: 10.1128/JVI.02236-09

Wang, S., Liu, H., Zu, X., Liu, Y., Chen, L., Zhu, X., et al. (2016). The ubiquitin-proteasome system is essential for the productive entry of Japanese encephalitis virus. *Virology* 498, 116–127. doi: 10.1016/j.virol.2016.08.013

Wang, Y., Jia, L., Shen, J., Wang, Y., Fu, Z., Su, S. A., et al. (2018). Cathepsin B aggravates coxsackievirus B3-induced myocarditis through activating the inflammasome and promoting pyroptosis. *PLoS Pathog.* 14:e1006872. doi: 10.1371/journal.ppat.1006872

Wei, S., Huang, Y., Huang, X., Cai, J., Yan, Y., Guo, C., et al. (2014). Characterization of cathepsin B gene from orange-spotted grouper, Epinephelus coioides involved in SGIV infection. *Fish Shellfish Immunol.* 36, 194–205. doi: 10.1016/j.fsi.2013.11.006

Wertz, I. E., and Dixit, V. M. (2008). Ubiquitin-mediated regulation of TNFR1 signaling. *Cytokine. Growth. Factor. Rev.* 19, 313–324. doi: 10.1016/j.cytogfr.2008.04.014

Widjaja, I., de Vries, E., Tscherne, D. M., García-Sastre, A., Rottier, P. J., and de Haan, C. A. (2010). Inhibition of the ubiquitin-proteasome system affects influenza A virus infection at a postfusion step. *J. Virol.* 84, 9625–9631. doi: 10.1128/JVI.01048-10

Yao, T., and Ndoja, A. (2012). Regulation of gene expression by the ubiquitin-proteasome system. *Semin. Cell Dev. Biol.* 23, 523–529. doi: 10.1016/j.semcdb.2012.02.006

Zhao, J., and Goldberg, A. L. (2016). Coordinate regulation of autophagy and the ubiquitin proteasome system by MTOR. *Autophagy* 12, 1967–1970. doi: 10.1080/15548627.2016.1205770

Zhao, L., Wang, H., Sun, X., and Ding, Y. (2010). Comparative proteomic analysis identifies proteins associated with the development and progression of colorectal carcinoma. *FEBS J.* 277, 4195–4204. doi: 10.1111/j.1742-4658.2010.07808.x

Zhao, Z., Ke, F., Huang, Y. H., Zhao, J. G., Gui, J. F., and Zhang, Q. Y. (2008). Identification and characterization of a novel envelope protein in *Rana grylio* virus. *J. Gen. Virol.* 89, 1866–1872. doi: 10.1099/vir.0.2008/000810-0

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 05 September 2018
doi: 10.3389/fmicb.2018.02132

Establishment of an Efficient and Flexible Genetic Manipulation Platform Based on a Fosmid Library for Rapid Generation of Recombinant Pseudorabies Virus

Mo Zhou, Muhammad Abid, Hang Yin, Hongxia Wu, Teshale Teklue, Hua-Ji Qiu and Yuan Sun**

State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin, China

OPEN ACCESS

Edited by:
Bernard La Scola,
Aix-Marseille Université, France

Reviewed by:
Akatsuki Saito,
Osaka University, Japan
Takayuki Murata,
Fujita Health University, Japan

***Correspondence:**
Hua-Ji Qiu
qiuhuaji@caas.cn;
huajiqiu@hvri.ac.cn
Yuan Sun
sunyuan@caas.cn

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

Received: *15 June 2018*
Accepted: *20 August 2018*
Published: *05 September 2018*

Citation:
Zhou M, Abid M, Yin H, Wu H, Teklue T, Qiu H-J and Sun Y (2018) Establishment of an Efficient and Flexible Genetic Manipulation Platform Based on a Fosmid Library for Rapid Generation of Recombinant Pseudorabies Virus. Front. Microbiol. 9:2132. doi: 10.3389/fmicb.2018.02132

Conventional genetic engineering of pseudorabies virus (PRV) is essentially based on homologous recombination or bacterial artificial chromosome. However, these techniques require multiple plaque purification, which is labor-intensive and time-consuming. The aim of the present study was to develop an efficient, direct, and flexible genetic manipulation platform for PRV. To this end, the PRV genomic DNA was extracted from purified PRV virions and sheared into approximately 30–45-kb DNA fragments. After end-blunting and phosphorylation, the DNA fragments were separated by pulsed-field gel electrophoresis, the recovered DNA fragments were inserted into the cloning-ready fosmids. The fosmids were then transformed into *Escherichia coli* and selected clones were end-sequenced for full-length genome assembly. Overlapping fosmid combinations that cover the complete genome of PRV were directly transfected into Vero cells and PRV was rescued. The morphology and one-step growth curve of the rescued virus were indistinguishable from those of the parent virus. Based on this system, a recombinant PRV expressing enhanced green fluorescent protein fused with the VP26 gene was generated within 2 weeks, and this recombinant virus can be used to observe the capsid transport in axons. The new genetic manipulation platform developed in the present study is an efficient, flexible, and stable method for the study of the PRV life cycle and development of novel vaccines.

Keywords: pseudorabies virus, fosmid library, full-length genome assemble, genetic manipulation platform, recombinant PRV

INTRODUCTION

Pseudorabies (PR), also known as Aujeszky's disease, is caused by pseudorabies virus (PRV) in the *Herpesviridae* family, which mainly affects swine and occasionally transmitted from pigs to cattle, sheep, goats, dogs, and cats (Masot et al., 2017; Zhou et al., 2017). Pigs are the only known natural reservoir for the virus. PRV has the ability to produce latent or clinically inapparent infections, which is transmitted between infected and non-infected pigs by nose-to-nose contact

(Smith and Enquist, 2000; Pomeranz et al., 2017). The mortality in piglets <1 month of age approaches to 100%. A virulent PRV variant has emerged and become prevalent in China since 2011. The disease caused by this PRV variant is characterized by neurological signs and high mortality among newborn piglets (Wang et al., 2016, 2017).

PRV is a linear, double-stranded DNA virus of about 143-kb and consists of unique long (UL) region, unique short (US), internal repeat short (IRS), and terminal repeat short (TRS) (Pomeranz et al., 2017). The genome contains at least 70 open reading frames (ORFs) that encode 70–100 viral proteins, including structural proteins, virulence-related proteins and replicase (Pomeranz et al., 2005; An et al., 2013). The marked progress in molecular biotechnology has significantly contributed to the study of other viruses' replication and vaccine. However, due to the large size of the PRV genome, the gene modification remains a difficult task. Earlier on, the recombinant PRV was generated by homologous recombination in permissive cells. Bacterial artificial chromosome (BAC) was used later and allowed cloning and manipulation of the whole genome in *Escherichia coli* (Lerma et al., 2016). The BAC system was more efficient than homologous recombination; however, from previous reports (Gu et al., 2015; Guo et al., 2016), it can be noticed that generation of recombinant BAC construct is time-consuming and labor-intensive due to several rounds of plaque purification and homologous recombination. However, generation of fosmid library is more efficient and minimizes the need of the above-mentioned steps.

The CopyControl cloning system of pCC1^{TM} Vector has a similar backbone to BAC and contains both a single-copy and the high-copy *ori*V origin of replication. Therefore, this system combines the clone stability afforded by single-copy cloning with the advantages of high DNA yields obtained by high-copy vector (Kim et al., 2003; Cunningham et al., 2006). Fosmids have been proved to have a high structural stability and found to maintain human DNA effectively even after 100 generations of bacterial growth, which was used for constructing stable libraries from complex genomes (Magrini et al., 2004; Zhang et al., 2007). A fosmid library is prepared by extracting the genomic DNA from the target organism, generating random genomic DNA fragments and cloning them into the fosmid vector (De Tomaso and Weissman, 2003; Moon and Magor, 2004). Therefore, construction of the infectious clones of large DNA viruses based on the fosmid library could alleviate the above drawbacks of BAC and allow manipulations of the viral genome more efficiently.

In this study, we constructed a fosmid library for the PRV-TJ strain, and an infectious progeny virus (rPRV-TJ) was rescued by directly transfecting fosmid sets into Vero cells. Moreover, a reporter virus (rPRV-VP26-EGFP) stably expressing enhanced green fluorescent protein (EGFP) was generated robustly via the Red/ET recombination. This study provides a foundation for rapid and accurate modification of the PRV genome. Meanwhile, the genetic manipulation of PRV based on the fosmid library also opens an exciting possibility and applicability for engineering other large DNA viruses in dissecting and probing genes of unknown functions.

MATERIALS AND METHODS

Virus Strain and Cells

The PRV-TJ strain (GenBank accession number: KJ789182.1) was isolated from a pig farm with a PR outbreak in Tianjin, China in 2012 and stored at −70°C and propagated in the porcine kidney 15 (PK-15) cell line (Luo et al., 2014). PK-15 and Vero cells were obtained from China Center for Type Culture Collection (CCTCC, Wuhan, China) and maintained at 37°C with 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) (ThermoFisher Scientific, Carlsbad, CA, United States) supplemented with 10% fetal bovine serum (Gibco, Grand Island, NY, United States). Dorsal root ganglions (DRGs) were isolated from newborn mice and cultured in Neurobasal medium (Gibco) supplemented with 100 ng/ml nerve growth factor 2.5S (Invitrogen), 2% B-27 (Gibco) and 1% penicillin and streptomycin with 2 mM glutamine (Invitrogen). The Animal Ethics Committee approval number is Heilongjiang-SYXK-2006-032. We conducted all the experiments in Biosafety Level II laboratory following strict biosecurity measures according to instructions of Harbin Veterinary Research Institute.

Extraction of High-Quality PRV Genomic DNA and Construction of a Fosmid Library Covering the Full-Length Genome of PRV

The PRV-TJ strain propagated in PK-15 cell line was used to isolate genome for fosmid library construction according to the method described previously (Smith and Enquist, 1999). In brief, 10 75 cm^2 flasks of confluent PK-15 cells were infected with the PRV-TJ strain at a multiplicity of infection (MOI) of 5. The cells were then incubated at 37°C for 15 h, and harvested by scraping. The scrapped cells were then washed twice with phosphate-buffered saline (PBS). The final cell pellet was re-suspended in 10 ml of LCM buffer (130 mM KCl, 30 mM Tris [pH 7.4], 5 mM $MgCl_2$, 0.5 mM EDTA, 0.5% nonidet P-40 [NP-40], and 0.043% 2-mercaptoethanol). The virus particle was extracted with Freon from re-suspended pellet and then the nucleocapsid pellets were extracted by centrifugation through two LCM buffer-based glycerol step gradients (8 ml of 5% glycerol and 16 ml of 45% glycerol) at 26,000 rpm for 2.5 h at 4°C. The nucleocapsid pellets were used to extract the genome. DNA quality was assessed by NanoDrop^{TM} 2000 (Thermo Scientific) and transfection. For transfection, monolayer of Vero cells grown on 6-well-plate was washed with PBS and then 2 ml of DMEM without antibiotics was added in each well for 1 h. The Vero cells were transfected with 2 μg of PRV genomic DNA using X-treme GENE HP DNA transfection reagent (Roche). The ratio of X-treme GENE HP DNA transfection reagent (μl) to PRV genomic DNA (μg) was 1:1.

Twenty microgram of PRV genomic DNA (at a concentration of 500 ng/μl) was pipetted 800 times with a 200-μl tip to shear the genomic DNA into approximately 30–45-kb fragments. To determine proper shearing, 1 μl of sheared DNA was analyzed on a 1% gold agarose gel by pulsed-field gel electrophoresis (PFGE)

using Fosmid Control DNA (Epicentre) and a Lambda DNA-Mono Cut Mix (New England BioLabs) as size marker. In order to generate 5′-phosphorylated DNA, 20 μg of sheared genomic DNA was end repaired using the End-Repairing Enzyme Mix according to the CopyControl™ Fosmid Library Production Kit. Following end repair, the genomic DNA was size selected on a low-melting point agarose gel by PFGE. DNA fragments ranging from 33 to 48-kb in size were excised from the gel and recovered using GELase (Epicentre) according to the instruction manual. The recovered fragments were then ligated into the pCC1FOS cloning-ready vector at room temperature for 4 h. The ligation mixture was subsequently packaged using MaxPlax Lambda Packaging Extracts. Ten microliter of the packaged phage was then added to 100 μl of EPI300-T1 cells. The infected EPI300-T1 cells were spread on the LB ager plate containing 12.5 μg/ml chloramphenicol.

Fosmid Sequencing and Full-Length Genome Assembly

The resulting number of clones for a complete fosmid library covering the entire PRV genome is 92. Therefore, 200 clones were randomly picked and cultured overnight in 5-ml LB liquid medium containing 12.5 μg/ml chloramphenicol and 50 μl of auto-induction solution (Epicentre). The fosmids were extracted using ZR BAC DNA Miniprep Kit (Zymo Research). Fosmid end-sequencing was performed using pCC1FOS sequencing primers. The inserted sequences of all the fosmids were screened by BLAST. All sequences with 100% identity were screened out from the data set. The fosmids that cover the complete PRV genome were used to assemble the full-length genome.

Rescue of the Recombinant PRV

Ten overlapping fosmid combinations (each group containing five fosmids) that cover the complete PRV genome were used for virus rescue. Briefly, 80–90% confluent Vero cells grown on 10-cm plates were washed with PBS and then cultured with 10 ml of DMEM medium without antibiotics for 1 h. Meanwhile, five overlapping fosmids in each group (2 μg each) were gently mixed with 30 μl of X-treme GENE HP DNA transfection reagent in 1 ml of DMEM and incubated at room temperature for 20 min. The mixture was added into the above-mentioned Vero cell monolayers and the transfected cells were incubated at 37°C with 5% CO_2. At 3 days post-transfection (dpt), the cell supernatant was harvested when most cells showed cytopathic effects (CPEs) for virus passaging and further characterization. Vero cells transfected with the fosmid sets missing one fosmid served as negative control.

Immunofluorescence Assay (IFA)

To confirm the rescued virus (rPRV-TJ), a swine anti-PRV serum derived from the PRV-TJ strain-infected pigs was used as primary antibody in indirect immunofluorescence analysis. PK-15 cells were seeded in 96-well plates and cultured in DMEM containing 5% FBS. The confluent cell monolayers were infected with serially 10-fold diluted rPRV-TJ for 36 h. The cells were fixed with ethanol for 30 min at −30°C, followed by incubation with swine anti-PRV sera (diluted 1:300 with PBS) for 2 h at 37°C and then with Alexa 488-conjugated goat anti-pig IgG (Thermo Fisher Scientific) (1:1,000) for 1 h at 37°C. Images were captured using an Olympus CK40 microscope.

PCR

To confirm the integrity of rPRV-TJ, the gB and gE genes were detected by PCR using the genome of rPRV-TJ as a template, the genome of PRV-TJ was used as the positive control. The specific primers for gE (5′-TGGCTCTGCGTGCTGTGCTC-3′ and 5′-CATTCGTCACTTCCGGTTTC-3′) and gB (5′-GGGGTTG GACAGGAAGGACACCA-3′ and 5′-AACCAGCTGCACGCT CAA-3′) were used. TaKaRa LA Taq™ with GC Buffer (TaKaRa) was used for PCR amplification. The reaction mixtures were performed in a final volume of 20 μl, containing 2 μl dNTP mixture, 1.0 μM concentration of each primer, 10 μl 2× GC Buffer I, 2 μl of virus DNA sample, and 0.25 μl of LA Taq. Reactions were conducted in an automated DNA thermal cycler (Bio-Rad, United States). The thermo-cycling condition was denaturation for 5 min at 95°C, followed by 35 cycles that each consisted of a denaturation step at 95°C for 30 s, an annealing step at 60°C for 30 s, and an extension at 72°C for 1 min, and the final extension at 72°C for 10 min.

Pulsed-Field Gel Electrophoresis

The genomic DNA (10 μg) of the rescued or parental virus was digested with *Kpn*I, *Nco*I, and *Pst*I, respectively, for 5 h at 37°C. The reaction was transferred to 70°C for 10 min to inactivate the restriction enzymes. The digested samples were analyzed by PFGE in a 1% (w/v) gold agarose gel in 0.5× TBE buffer at 6 V/cm and 14°C for 16 h. The λ DNA-Mono Cut Mix was used as standard.

Electron Microscopy

Vero cells were infected with rPRV-TJ and PRV-TJ and harvested at 48 h post-infection (hpi). Cell culture medium was centrifuged at 3,000 × *g* for 10 min, the supernatant was collected and centrifuged at 10,000 × *g* for 10 min, and then the pellet was resuspended in PBS. The sample was negatively stained with 2% phosphotungstic acid, the morphology of the rescued virus was observed under electron microscope and PRV-TJ particles as positive control to compare the morphology.

Plaque Assay

rPRV-TJ and PRV-TJ were serially 10-fold diluted in DMEM. One hundred microliter of diluted sample was inoculated onto Vero cell monolayers in 12-well culture plates. After incubation for 1 h at 37°C, the monolayers were washed twice with DMEM and overlaid with 2 ml of DMEM containing 1% low melting point agarose. The plaque-forming units (PFUs) were determined at 5 days post-infection.

Replication Kinetics of the Rescued PRV

The virus titers of rPRV-TJ and PRV-TJ were determined according to the Reed-Muench method. PK-15 cells cultured in a

24-well plate were infected with rPRV-TJ and PRV-TJ at an MOI of 10 and incubated on ice for 1 h. Thereafter, the inoculum was replaced with pre-warmed fresh medium and cells were further incubated for 1 h at 37°C and rinsed for 2 min with citrate buffer (pH 3.0) to inactivate any unabsorbed virus. Then fresh medium was added and the cells were incubated at 37°C in 5% CO_2. The cultures were harvested at 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, 48, 60, and 72 hpi. The titers of all the collected samples were determined in duplicates on monolayers of PK-15 cells, and the average of each was calculated as described previously (Luo et al., 2014).

Generation of a Recombinant PRV Using the Fosmid Library and Red/ET System

Capsid assembly occurs in the nucleus of infected cells, initially with a spherical pro-capsid precursor built around a protein scaffold that matures into a DNA-containing capsid. VP26 is one of the first herpesvirus proteins to be fused with a fluorescent protein. Capsid-tagged virus mutants have been used to study capsid transport, intra-nuclear capsid dynamics, and nuclear egress (Desai and Person, 1998; Hogue et al., 2015). Therefore, in this study, the EGFP gene was inserted between the second and third codons of the VP26 gene by Counter Selection BAC Modification Kit (Gene Bridges, Berkeley, CA, United States) according to the manufacturer's instructions. In the first step, the Red/ET expression plasmid (pRed/ET) and the fosmid were co-transformed into competent *E. coli* DH10B cells by electroporation. In the next step, the antibiotic selectable cassette (*rpsL-neo*) flanked by the homology arms was generated by PCR amplification with specific primers in **Table 1** and inserted into the target site of the fosmid by the Red/ET-mediated recombination. To fuse the EGFP gene with the VP26 gene, the electro-component cells were prepared from the cells containing modified fosmid carrying a *rpsL-neo* cassette. In advance, the linear DNA fragment of the EGFP gene with homology arms was amplified with specific primers in **Table 1**. The EGFP gene flanked by two oligonucleotide homology arms was transformed into the prepared electro-component cells to replace the *rpsL-neo* cassette by the Red/ET-mediated recombination. The modified VP26 ORFs were amplified and sequenced. Finally, the modified fosmid plus the other fosmids were transfected into Vero cells to rescue the virus. Vero cells transfected with the fosmid set missing one fosmid served as a negative control and those transfected with the unmodified fosmid set as a positive control, and the recombinant virus expressing the VP26-EGFP fusion protein was rescued.

TABLE 1 | Primers for Red/ET recombination.

Names	Sequences (5′–3′)	Target genes	Insert position
F-VP26-rpsL	GCGCGCGGGGGCGCGCAC AGACGCGCGCTCCCCGCCG AGCCATCATGTCC*GGCCTGG TGATGATGGCGGGATCG*	rpsL	After the second codon of VP26
R-VP26-rpsL	GCGCCCTCGAGCGTCTGC GCGGTGATCGTCCGGGGAT TGTTCGGGTCGAA*TCAGAA GAACTCGTCAAGAAGGCG*		
F-VP26-EGFP	GCGCGCGGGGGCGCGCA CAGACGCGCGCTCCCCGC CGAGCCATCATGTCC*GTGA GCAAGGGCGAGGAGC*	EGFP	After the second codon of VP26
R-VP26-EGFP	GCGCCCTCGAGCGTCTGC GCGGTGATCGTCCGGGGA TTGTTCGGGTCGAA*TCTAG ATCCGGTGGATCCCG*		
F-identify rpsL	CATCATCCTGAACATGCG		
R-identify rpsL	GCTGCTGTAGTCGCTGGTG		

Characterization of the Recombinant Virus

To evaluate the genetic stability of the reporter virus containing the EGFP gene, rPRV-VP26-EGFP was passaged in PK-15 cells for 20 generations. The essential genes (gB and gE) were amplified using the genome of the rescued virus as a template according to the above-mentioned method, and the inserted gene (EGFP) was also amplified with the specific primers (5′-CATCATCCTGAACATGCG-3′ and 5′-CAT CATCCTGAACATGCG-3′). Replication kinetics and PFU of the rescued PRV were analyzed according to the above mentioned method.

Infection of Neurons With rPRV-VP26-EGFP

Microfluidic device is a useful tool for neuroscience research, which can separate the neuron and axons (Taylor et al., 2003; Harris et al., 2007). In this study, the neuron microfluidic device was used to observe the capsid transportation between neurons and axons. The microfluidic device can separate the soma and axonal side of DRGs. The DRGs were isolated from 2-day neonatal BALB/C mice and loaded into the axonal soma of the devices as described in previous report (Harris et al., 2007). One day after seeding, 5 mM arabinofuranoside (AraC; Sigma-Aldrich) was added for at least 2 days to eliminate non-neuronal cells. Neurons were cultured for around 5 days, the axons grown and flown through to the axon side. The rPRV-VP26-EGFP was infected to the soma side at an MOI of 5, and thus, the axon side could not contact rPRV-VP26-EGFP. Therefore, we excluded the possibility that the virus entered both neurons and axons. The green fluorescent capsids were observed at 12 hpi under a fluorescence microscope.

RESULTS

Extraction of High-Quality PRV Genomic DNA

The concentration of PRV genomic DNA was determined by Thermo Scientific NanoDrop™ 2000. The concentration was 692.5 ng/μl and A260/A280 was 1.80. The full-length genome of

TABLE 2 | Fosmids that cover the entire genome of PRV.

Fosmid	Location in genome (nt)	Size (bp)	Fosmid	Location in genome (nt)	Size (bp)
a	1-41633	41,633	k	66341-111030	4,4690
b	1-35363	35,363	l	49114-89118	4,0005
c	1-32129	32,129	m	64783-95747	3,0947
d	1-29834	37,955	n	87260-116855	2,9596
e	29544-67498	32,055	o	62988-100570	3,7583
f	37942-69996	35,778	p	82459-124194	4,1736
g	23721-59498	31,544	q	81316-117252	3,5937
h	55457-87000	34,390	r	86680-125105	3,8426
i	48838-83227	34,390	s	113582-143642	3,00061
j	66874-98904	32,031			

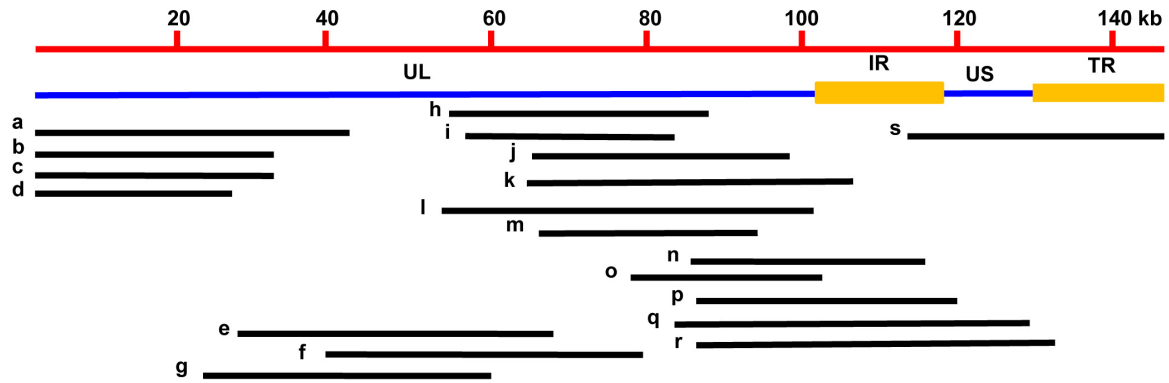


FIGURE 1 | Fosmid library construction for the PRV-TJ strain. Genomic structure of PRV and the five fosmid DNAs used for rescue of PRV. Numbers show the location of each fosmid fragment in the PRV genome.

PRV was used to transfect Vero cells using X-treme GENE HP DNA transfection reagent. At 24 h post-transfection, CPEs were observed in most transfected cells. The cell culture supernatant was harvested and used to inoculate PK-15 cells and the CPE became obvious (data not shown). Therefore, the concentration and quality of genomic DNA could thus satisfy the needs of fosmid library construction.

Generation of the Fosmid Library for PRV

A total of 200 clones were randomly picked from the fosmid library for end-sequencing, representing more than twofold coverage of PRV genome. A total of 180 clones contained DNA fragments of PRV-TJ, a majority of which contains inserts of 30–40-kb. A fosmid library covering the complete genome of PRV was established. Nineteen fosmids were selected for generating the fosmid-combinations that cover the entire genome of PRV (**Table 2**). Ten sets of overlapping fosmid-combinations were prepared to rescue the recombinant PRV, each consisting of five overlapping fosmids (**Figure 1** and **Table 3**).

Rescue of PRV From Overlapping Fosmids and Characterization of the Rescued PRV

Ten sets of fosmid were transfected into Vero cells to rescue the virus. The concentration of each fosmid was determined in ng/μl and then volume was adjusted according to 2 μg for each fosmid. Each group contain five overlapping fosmids, 2 μg of each fosmid was gently mixed with 30 μl of X-treme GENE HP DNA transfection reagent for transfection. CPEs were observed in Vero cells at 2 dpt from sets 1, 3, 4, 5, and 6, but not from other sets and the negative control. The PK-15 cells infected with rPRV-TJ were assayed by IFA using swine anti-PRV sera at 24 hpi (**Figure 2A**). The expected bands of the gB and gE genes were amplified from the genomic DNA of rPRV-TJ (**Figure 2B**). Under electron microscope, the rPRV-TJ particles showed similar morphology to that of the parental virus with an apparently external envelope (**Figure 2C**). The genome of rPRV-TJ was digested with *Kpn*I, *Nco*I and *Pst*I, and analyzed by PFGE. The digestion patterns of *Kpn*I and *Nco*I are 100%, according to our observations, the *Pst*I digestion pattern

TABLE 3 | Fosmid combinations that cover the entire genome of PRV.

Group	Combinations	CPE	Group	Combinations	CPE
1	b + e + h+p+s	+	6	b + e + i+q+s	+
2	a + f + k+r+s	–	7	a + f + j+q+s	–
3	c + g + l+r+s	+	8	a + f + m+n+s	+
4	a + f + o+q+s	+	9	d + g + h+p+s	–
5	b + e + h+q+s	+	10	d + g + l+q+s	–

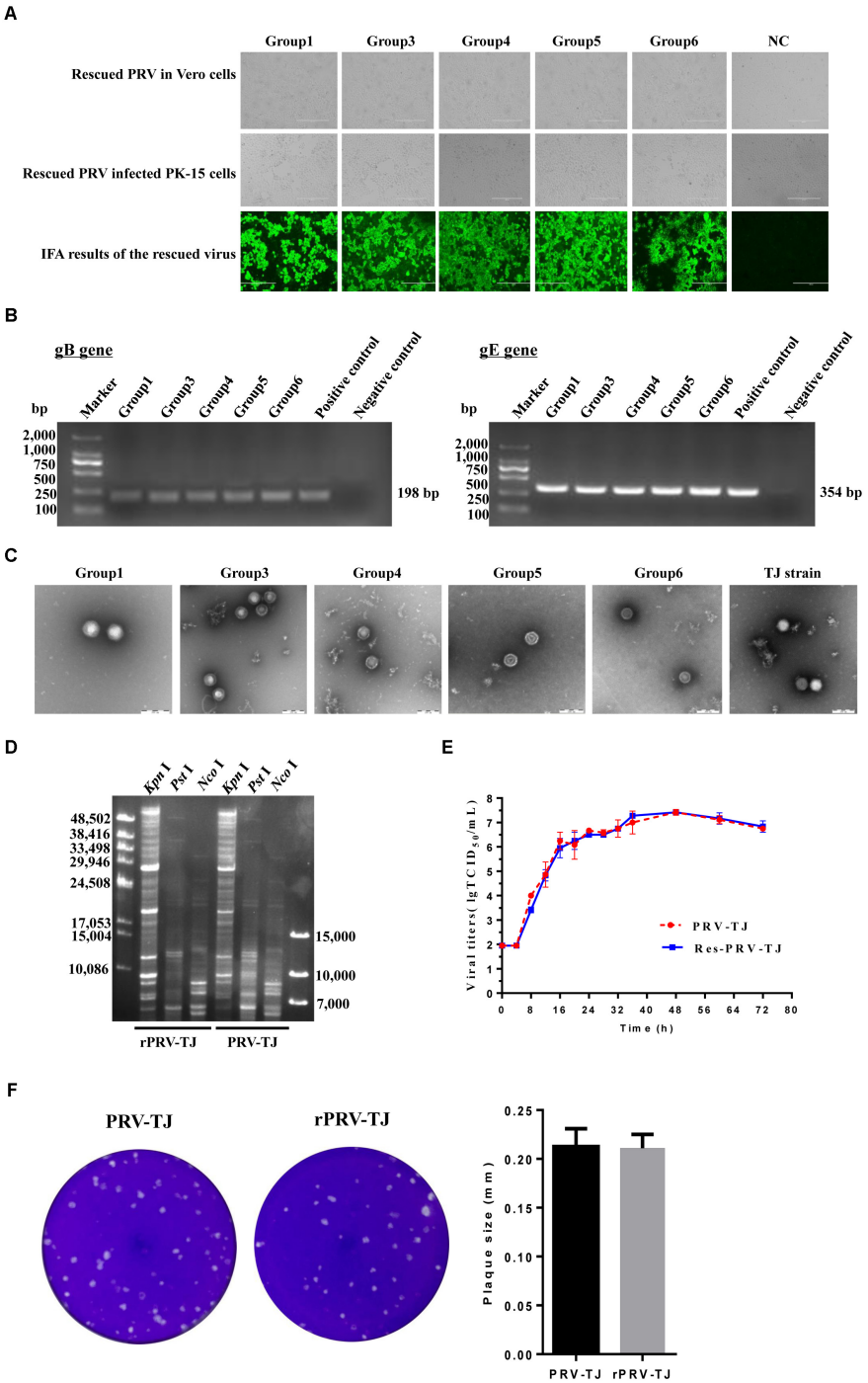


FIGURE 2 | Identification of the rescued PRV. **(A)** Transfection of the Vero cells with the fosmid combinations, the cytopathic effects (CPEs) was observed. Vero cells transfected with the fosmid set missing one fosmid was used as negative control. The culture supernatants of transfected cells was collected and used to infect PK-15 cells, the CPEs were also observed. The rescued PRV was detected by IFA with an anti-PRV serum. **(B)** PCR amplification of the gB and gE genes. The gB and gE genes were amplified using the genome of the rescued virus as a template. The genome of the parental PRV was used as a positive control. The irrelevant genome was used as a negative control. **(C)** Transmission electron of viral DNA of the rescued and the parent PRV. PRV-TJ particles were used as positive control. Scale bars are presented. **(D)** The restriction profiles of the rescued and the parent PRV in 1% agarose. The genome of the rescued and the parent PRV was digested with *Kpn*I, *Nco*I, and *Pst*I. **(E)** One-step growth curves of the rescued and the parent PRV. PK-15 cells cultured in 24-well plates were infected with the rescued or the parent PRV at a multiplicity of infection (MOI) of 10, after incubated on ice and rinsed with citrate buffer, the virus was harvested from both the medium and cells at 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, 48, 60, and 72 h post-infection (hpi), and titers were determined. Titration was performed in triplicates; error bars represent standard errors of the mean. **(F)** The plaque size of the rescued and the parent PRV. The rescued and the parent PRV were 10-fold serial diluted at a titer at which a single plaque could be formed. The experiments were performed in triplicates, and the representative results were shown.

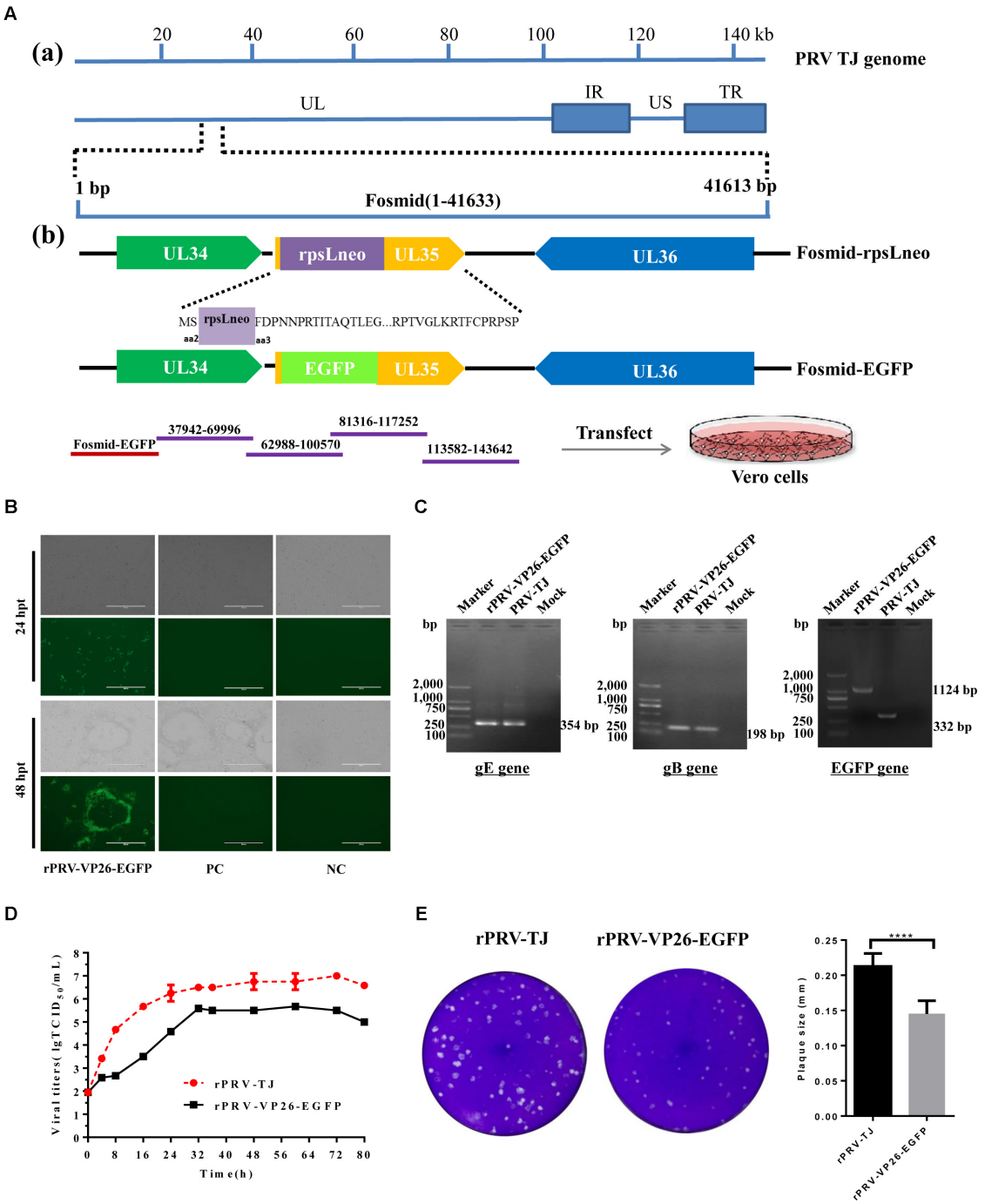


FIGURE 3 | Generation of rPRV-VP26-EGFP. **(A)** Schematic representation of Fosmid(1-41,633) modification for the generation of a recombinant PRV-TJ strain expressing the EGFP gene. *a*: The flow diagram of the PRV genome and the position of Fosmid(1-41,633). *b*: Construction procedure of the intermediate fosmids and Fosmid-EGFP. For fosmid modification, the antibiotics-selectable cassette (*rpsL-neo*) flanked by two oligonucleotide homology arms was inserted between the second and third amino acids of VP26 by the Red/ET-mediated recombination. The EGFP gene flanked by the same homology arms was used to replace the *rpsL-neo* cassette to generate Fosmid-EGFP. For whole genome assembly, Fosmid-EGFP and other fosmids in Group 4 were transfected into Vero cells to generate a recombinant PRV expressing the EGFP gene. **(B)** Fosmid-EGFP and other fosmids in Group 4 were transfected into Vero cells to rescue the virus. Vero cells that were transfected with the fosmid set missing one fosmid was served as negative control and Vero cells transfected with the un-modified fosmid set as positive control. The images were taken at 24 and 72 hpi. **(C)** PCR amplification of the gB, gE, and EGFP genes. The gB, gE, and EGFP genes were amplified using the genome of the rescued virus as a template. The genome of the parental virus was used as a positive control. The externally located primers, which across the insertion site, were used to amplify the EGFP gene. **(D)** The replication kinetics of rPRV-TJ and rPRV-VP26-EGFP. The virus titers were calculated at 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, 48, 60, and 72 hpi. **(E)** Plaque size of rPRV-TJ and rPRV-VP26-EGFP. The virus was 10-fold diluted at a titer at which single plaques could be formed. The experiments were performed in triplicates, and representative results were shown.

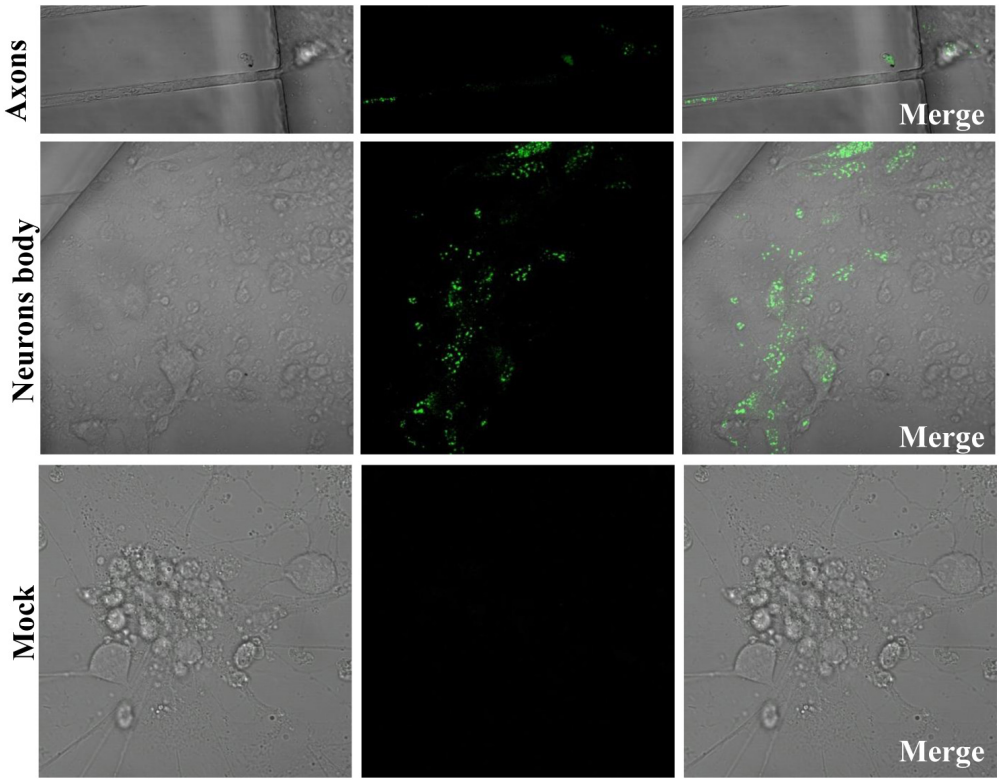


FIGURE 4 | rPRV-VP26-EGFP infection in neuron cells. Dorsal root ganglions (DRGs) were plated in the main channel of Microfluidics. After 5 days, DRGs were infected with rPRV-VP26-EGFP at an MOI of 5. Fluorescent images were taken 12 hpi, the green fluorescent capsids accumulated in cell bodies and transport between neuron bodies and axons.

is also similar and the difference may be due to band intensity (**Figure 2D**). The replication kinetics and plaque morphology of rPRV-TJ and the parental virus had no significant difference (**Figures 2E,F**).

Generation of rPRV-TJ Expressing EGFP

VP26 is the small capsid protein expressed early during replication, tagging VP26 is helpful to visualize alpha herpesvirus capsids under fluorescence microscopy (Hogue et al., 2015). Therefore, the VP26 gene was used for tagging the capsid protein by fusion with the coding sequence of EGFP in this study. To rescue rPRV-VP26-EGFP, we modified Fosmid(1-41,633) that harbors the VP26 gene (**Figure 3A**). The resulting modified fosmid and other fosmids in the combination were transfected into Vero cells. After 2 days CPEs and green fluorescence were detected (**Figure 3B**). CPEs were observed in the Vero cells transfected with the complete fosmid set but not with the set with one fosmid missing. Expected bands of the gB, gE, and EGFP genes were amplified from the genomic DNA of the rescued rPRV-VP26-EGFP (**Figure 3C**). The genomic DNA of PRV-TJ was used as the positive control. The externally located primers that cross the insertion site were used to amplify the EGFP gene. The growth kinetics of rPRV-VP26-EGFP was delayed compared with rPRV-TJ, the rPRV-VP26-EGFP has around 1 log defect in the peak virus titer relative to rPRV-TJ (**Figure 3D**). The plaque size of the rescued virus in Vero cells was also smaller than that of rPRV-TJ (**Figure 3E**).

Visualization of the EGFP-Tagged Capsids in Neuron Bodies and Axons

We generated a reporter PRV virus containing EGFP to monitor virus moving both toward the cell body (retrograde) and away from the cell body (anterograde), and thus can facilitate to study virus transport, intra-nuclear capsid dynamics and nuclear egress. Therefore, using the EGFP fused PRV-TJ VP26 we were able to analyze the virus transport in axons. The soma side was infected with the rPRV-VP26-EGFP, and the EGFP signals were imaged at 12 hpi, the EGFP-tagged capsid transported from the soma side to the axons that in the connecting region of soma and axonal side, there was no EGFP signal in the mock cells (**Figure 4**). The results indicate that rPRV-VP26-EGFP allows visualization of the EGFP-tagged capsid transport between neuron bodies and axons.

DISCUSSION

Generating recombinant PRV using traditional methods, such as plasmid transfection plus virus co-infection, often are inefficient,

labor-intensive, and time-consuming due to the need of cloning and purification processes. In addition, the insertion of a selection marker is another tedious and time-consuming process. Infectious BACs of herpesviruses are powerful tools for genetic manipulation (Tobler and Fraefel, 2015; Close et al., 2017). However, construction of BAC clones usually takes several months and the presence of BAC vector sequence in the viral genomes often causes genetic and phenotypic alterations (Zhao et al., 2008; Zhou et al., 2010). Alternatively, fosmid library provides a powerful platform for rescue of viruses in recent studies (Liu et al., 2011, 2016; Li et al., 2016). This system allows unbiased inclusion of only viral genomic DNA fragments and seamless cloning without any genetic scar in comparison to fragments generated by restriction enzymes. Therefore, in order to improve genetic manipulation platform for PRV, in this study, a PRV fosmid library was constructed. High-quality PRV genomic DNA preparation is a critical step to construct a fosmid library. Preparation of high molecular weight DNA (around 40-kb) is also an important step as well as the basis for constructing a high quality library. Therefore, the quality of PRV genomic DNA was assessed by transfection to make sure the integrity of the PRV genomic DNA for the fosmid library, after that the genomic DNA was sheared into approximately 30–45-kb fragments, separated by PFGE, excised from gel and recovered. Thus, the average insert size for the fosmid library is 30–40-kb, and the number of clones with an insert is >90% and a high quality fosmid library was generated in this study.

Different sets of fosmids that cover the complete genome of PRV were used to rescue recombinant PRVs. Ten overlapping fosmid combinations were used to rescue virus. However, some combinations were successful to rescue virus, whereas others did not produce CPEs at 3 dpt. The overlapping region of each combination was different, which may be the reason of the variable efficiency among these combinations to rescue the virus. The aim of this study was to screen the fosmid combinations that could produce CPEs in short time period, only sets 1, 3, 4, 5, and 6 could rescue the virus within 3 dpt, other combinations were also monitored up to 5 days and no CPEs were observed, so we did not proceed to further steps.

The fosmid library-based genetic manipulation platform for PRV offers several advantages over the conventional technology. First, PRV genome was randomly fragmented into suitable 40-kb DNA fragments for cloning into the pCC1FOS vector is much simpler and far less time-consuming because there is no need to isolate high molecular weight DNA, or perform partial restriction enzyme digestions. Furthermore, high cloning efficiency of fosmid also makes it easy to achieve full genome coverage. Therefore, fosmid library facilitates the generation of infectious PRV. Second, fosmid vector maintains the clones as single copy, thereby enhancing insert stability. Meanwhile, the fosmid vector contains an inducible, high-copy origin oriV, which increases copy number for higher yields in the presence of an inducer without compromising insert stability. Third, the established methodology is flexible to rescue recombinant viruses from overlapping fragments of cloned viral DNA, which is based on minimal sequence modification in bacteria and allows the modification of any essential genes of PRV by Red/ET recombination. The modification of individual 40-kb-fosmid was more convenient than the oversized BAC DNA constructs. Fourth, the construction of homology arms and plaque purification are not required in this system. Therefore, the use of the fosmid library greatly reduces the time and labor for generating recombinant PRVs.

In this study, we constructed a fosmid library of PRV and rescued the virus by transfecting overlapping fosmids into Vero cells. The typical virus biological characteristics such as morphology and one-step growth curve analyses revealed that the rescued virus was indistinguishable from the parental virus. The recombinant virus expressing EGFP fused to VP26 was generated based on the fosmid library-based genetic manipulation platform, which allows further monitoring the pathogenicity of PRV *in vivo* and *in vitro*. Furthermore, rPRV-VP26-EGFP allows us to monitor visually the localization of the virus at various stages of infection. We found that rPRV-VP26-EGFP caused about 10-fold defect in single-step virus replication (**Figure 3D**). Some studies also reported the effect of fluorescent protein fusions to VP26 on virus replication kinetics, cell-to-cell spread and pathogenesis *in vivo* (Krautwald et al., 2008; Hogue et al., 2015). The possible explanation to these differences might be the insertion of EGFP gene affect capsid assembly. In addition, the plaque sizes of the recombinants were smaller than those of PRV-TJ (**Figure 3E**). Some reports indicated that fluorescent protein fuses to VP26 affects cell-to-cell spread of the recombinant virus. The smaller plaque size may indicate the cell-to-cell spread ability of the recombinant virus become lower. Therefore, the smaller plaque size may correlate with lower replication capability of rPRV-VP26-EGFP.

In summary, this genetic manipulation platform provides an opportunity to explore the biology of PRV in depth. Similarly, the method of fosmid constructing platform can be extended to other large double-stranded DNA viruses. This platform will be directly used for the development of novel bivalent, trivalent and marker vaccines. We believe that any newly emerged PRV strain and other DNA viruses can be manipulated using this platform, and possibly vaccines could be developed in a short time period.

AUTHOR CONTRIBUTIONS

MZ, YS, and H-JQ designed the study. MZ wrote the manuscript. MZ, MA, HY, HW, and TT performed the experiments. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (No. 2016YFD0500105), the National Natural Sciences Foundation of China (Nos. 31570149 and 31802163), the China Postdoctoral Science Foundation (No. 2017M620981), and the Heilongjiang Natural Sciences Foundation (No. QC2018029).

REFERENCES

An, T. Q., Peng, J. M., Tian, Z. J., Zhao, H. Y., Li, N., Liu, Y. M., et al. (2013). Pseudorabies virus variant in bartha-K61-vaccinated pigs, China. *Emerg. Infect. Dis.* 19, 1749–1755. doi: 10.3201/eid1911.130177

Close, W. L., Bhandari, A., Hojeij, M., and Pellett, P. E. (2017). Generation of a novel human cytomegalovirus bacterial artificial chromosome tailored for transduction of exogenous sequences. *Virus Res.* 242, 66–78. doi: 10.1016/j.virusres.2017.09.007

Cunningham, C., Hikima, J., Jenny, M. J., Chapman, R. W., Fang, G. C., Saski, C., et al. (2006). New resources for marine genomics: bacterial artificial chromosome libraries for the eastern and Pacific oysters (*Crassostrea virginica* and *C. gigas). Mar. Biotechnol.* 8, 521–533. doi: 10.1007/s10126-006-6013-9

Desai, P., and Person, S. (1998). Incorporation of the green fluorescent protein into the herpes simplex virus type 1 capsid. *J. Virol.* 72, 7563–7568.

De Tomaso, A. W., and Weissman, I. L. (2003). Construction and characterization of large-insert genomic libraries (BAC and fosmid) from the ascidian *Botryllus schlosseri* and initial physical mapping of a histocompatibility locus. *Mar. Biotechnol.* 5, 103–115. doi: 10.1007/s10126-002-0071-1

Gu, Z., Dong, J., Wang, J., Hou, C., Sun, H., Yang, W., et al. (2015). A novel inactivated gE/gI deleted pseudorabies virus (PRV) vaccine completely protects pigs from an emerged variant PRV challenge. *Virus Res.* 195, 57–63. doi: 10.1016/j.virusres.2014.09.003

Guo, J. C., Tang, Y. D., Zhao, K., Wang, T. Y., Liu, J. T., Gao, J. C., et al. (2016). Highly efficient CRISPR/Cas9-mediated homologous recombination promotes the rapid generation of bacterial artificial chromosomes of pseudorabies virus. *Front. Microbiol.* 7:2110. doi: 10.3389/fmicb.2016.02110

Harris, J., Lee, H., Vahidi, B., Tu, C., Cribbs, D., Cotman, C., et al. (2007). Non-plasma bonding of PDMS for inexpensive fabrication of microfluidic devices. *J. Vis. Exp.* 9:410. doi: 10.3791/410

Hogue, I. B., Bosse, J. B., Engel, E. A., Scherer, J., Hu, J. R., Del Rio, T., and Enquist, L. W. (2015). Fluorescent protein approaches in alpha herpesvirus research. *Viruses* 7, 5933–5961. doi: 10.3390/v7112915

Kim, C. G., Fujiyama, A., and Saitou, N. (2003). Construction of a gorilla fosmid library and its PCR screening system. *Genomics* 82, 571–575. doi: 10.1016/S0888-7543(03)00174-5

Krautwald, M., Maresch, C., Klupp, B. G., Fuchs, W., and Mettenleiter, T. C. (2008). Deletion or green fluorescent protein tagging of the pUL35 capsid component of pseudorabies virus impairs virus replication in cell culture and neuroinvasion in mice. *J. Gen. Virol.* 89, 1346–1351. doi: 10.1099/vir.0.83652-0

Lerma, L., Muñoz, A. L., Wagner, S., Dinu, M., Martín, B., and Tabarés, E. (2016). Construction of recombinant pseudorabies viruses by using PRV BACs deficient in IE180 or pac sequences: application of vBAC90D recombinant virus to production of PRV amplicons. *Virus Res.* 213, 274–282. doi: 10.1016/j.virusres.2015.11.028

Li, K., Liu, Y., Liu, C., Gao, L., Zhang, Y., Cui, H., et al. (2016). Recombinant marek's disease virus type 1 provides full protection against very virulent marek's and infectious bursal disease viruses in chickens. *Sci. Rep.* 6:39263. doi: 10.1038/srep39263

Liu, J., Chen, P., Jiang, Y., Wu, L., Zeng, X., Tian, G., et al. (2011). A duck enteritis virus-vectored bivalent live vaccine provides fast and complete protection against H5N1 avian influenza virus infection in ducks. *J. Virol.* 85, 10989–10998. doi: 10.1128/JVI.05420-11

Liu, Y., Li, K., Gao, Y., Gao, L., Zhong, L., Zhang, Y., et al. (2016). Recombinant marek's disease virus as a vector-based vaccine against Avian leukosis virus subgroup J in chicken. *Viruses* 8:301. doi: 10.3390/v8110301

Luo, Y., Li, N., Cong, X., Wang, C. H., Du, M., Li, L., et al. (2014). Pathogenicity and genomic characterization of a pseudorabies virus variant isolated from bartha-K61-vaccinated swine population in China. *Vet. Microbiol.* 174, 107–115. doi: 10.1016/j.vetmic.2014.09.003

Magrini, V., Warren, W. C., Wallis, J., Goldman, W. E., Xu, J., Mardis, E. R., et al. (2004). Fosmid-based physical mapping of the *Histoplasma capsulatum* genome. *Genome Res.* 14, 1603–1609. doi: 10.1101/gr.2361404

Masot, A. J., Gil, M., Risco, D., Jiménez, O. M., Núñez, J. I., and Redondo, E. (2017). Pseudorabies virus infection (Aujeszky's disease) in an Iberian lynx (*Lynx pardinus*) in Spain: a case report. *BMC Vet. Res.* 13:6. doi: 10.1186/s12917-016-0938-7

Moon, D. A., and Magor, K. E. (2004). Construction and characterization of a fosmid library for comparative analysis of the duck genome. *Anim. Genet.* 35, 417–418. doi: 10.1111/j.1365-2052.2004.01177.x

Pomeranz, L. E., Ekstrand, M. I., Latcha, K. N., Smith, G. A., Enquist, L. W., and Friedman, J. M. (2017). Gene expression profiling with cre-conditional pseudorabies virus reveals a subset of midbrain neurons that participate in reward circuitry. *J. Neurosci.* 37, 4128–4144. doi: 10.1523/JNEUROSCI.3193-16.2017

Pomeranz, L. E., Reynolds, A. E., and Hengartner, C. J. (2005). Molecular biology of pseudorabies virus: impact on neurovirology and veterinary medicine. *Microbiol. Mol. Biol. Rev.* 69, 462–500. doi: 10.1128/MMBR.69.3.462-500.2005

Smith, G. A., and Enquist, L. W. (1999). Construction and transposon mutagenesis in *Escherichia coli* of a full-length infectious clone of pseudorabies virus, an alphaherpesvirus. *J. Virol.* 73, 6405–6414.

Smith, G. A., and Enquist, L. W. (2000). A self-recombining bacterial artificial chromosome and its application for analysis of herpesvirus pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4873–4878. doi: 10.1073/pnas.080502497

Taylor, A. M., Rhee, S. W., Tu, C. H., Cribbs, D. H., Cotman, C. W., and Jeon, N. L. (2003). Microfluidic multicompartment device for neuroscience research. *Langmuir* 19, 1551–1556. doi: 10.1021/la026417v

Tobler, K., and Fraefel, C. (2015). Infectious delivery of alphaherpesvirus bacterial artificial chromosomes. *Methods Mol. Biol.* 1227, 217–230. doi: 10.1007/978-1-4939-1652-8_10

Wang, J., Guo, R., Qiao, Y., Xu, M., Wang, Z., Liu, Y., et al. (2016). An inactivated gE-deleted pseudorabies vaccine provides complete clinical protection and reduces virus shedding against challenge by a chinese pseudorabies variant. *BMC Vet. Res.* 12:277. doi: 10.1186/s12917-016-0897-z

Wang, X., Wu, C. X., Song, X. R., Chen, H. C., and Liu, Z. F. (2017). Comparison of pseudorabies virus China reference strain with emerging variants reveals independent virus evolution within specific geographic regions. *Virology* 506, 92–98. doi: 10.1016/j.virol.2017.03.013

Zhang, L., Bao, Z., Cheng, J., Li, H., Huang, X., Wang, S., et al. (2007). Fosmid library construction and initial analysis of end sequences in Zhikong scallop (*Chlamys farreri*). *Mar. Biotechnol.* 9, 606–612. doi: 10.1007/s10126-007-9014-4

Zhao, Y., Petherbridge, L., Smith, L. P., Baigent, S., and Nair, V. (2008). Self-excision of the BAC sequences from the recombinant marek's disease virus genome increases replication and pathogenicity. *Virol. J.* 5:19. doi: 10.1186/1743-422X-5-19

Zhou, F., Li, Q., Wong, S. W., and Gao, S. J. (2010). Autoexcision of bacterial artificial chromosome facilitated by terminal repeat-mediated homologous recombination: a novel approach for generating traceless genetic mutants of herpesviruses. *J. Virol.* 84, 2871–2880. doi: 10.1128/JVI.01734-09

Zhou, J., Li, S., Wang, X., Zou, M., and Gao, S. (2017). Bartha-k61 vaccine protects growing pigs against challenge with an emerging variant pseudorabies virus. *Vaccine* 35, 1161–1166. doi: 10.1016/j.vaccine.2017.01.003

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 April 2018
doi: 10.3389/fmicb.2018.00653

Antiviral Immunotoxin Against *Bovine herpesvirus*-1: Targeted Inhibition of Viral Replication and Apoptosis of Infected Cell

Jian Xu[1,2†], Xiaoyang Li[1,2,3†], Bo Jiang[1,2†], Xiaoyu Feng[4], Jing Wu[1,2,3], Yunhong Cai[1,2], Xixi Zhang[1,2], Xiufen Huang[1,2], Joshua E. Sealy[5], Munir Iqbal[5] and Yongqing Li[1,2]*

[1] Beijing Key Laboratory for Prevention and Control of Infectious Diseases in Livestock and Poultry, Beijing, China, [2] Institute of Animal Husbandry and Veterinary Medicine, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [3] College of Animal Science and Technology, Jiangxi Agricultural University, Nanchang, China, [4] Beijing Center for Animal Disease Control and Prevention, Beijing, China, [5] The Pirbright Institute, Woking, United Kingdom

OPEN ACCESS

Edited by:
Jonatas Abrahao,
Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Bruno Fernandes Mota,
Universidade Federal de Minas Gerais, Brazil
Danilo Oliveira,
Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brazil

***Correspondence:**
Yongqing Li
liyongqing@iasbaafs.net.cn;
chunyudady@sina.com

[†] These authors have contributed equally to this work.

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

Received: 07 February 2018
Accepted: 20 March 2018
Published: 04 April 2018

Citation:
Xu J, Li X, Jiang B, Feng X, Wu J, Cai Y, Zhang X, Huang X, Sealy JE, Iqbal M and Li Y (2018) Antiviral Immunotoxin Against Bovine herpesvirus-1: Targeted Inhibition of Viral Replication and Apoptosis of Infected Cell. Front. Microbiol. 9:653. doi: 10.3389/fmicb.2018.00653

Bovine herpesvirus 1 (BoHV-1) is a highly contagious viral pathogen which causes infectious bovine rhinotracheitis in cattle worldwide. Currently, there is no antiviral prophylactic treatment available capable of mitigating the disease impact and facilitating recovery from latent infection. In this study, we have engineered a novel recombinant anti-BoHV-1 immunotoxin construct termed "BoScFv-PE38" that consists of a single-chain monoclonal antibody fragment (scFv) fused with an active domain of *Pseudomonas* exotoxin A as a toxic effector (PE38). The recombinant BoScFv-PE38 immunotoxin expressed in a prokaryotic expression system has specific binding affinity for BoHV-1 glycoprotein D (gD) with a dissociation constant (Kd) of 12.81 nM and for BoHV-1 virus particles with a Kd value of 97.63 nM. We demonstrate that the recombinant BoScFv-PE38 is internalized into MDBK cell compartments that inhibit BoHV-1 replication with a half-maximal inhibitory concentration (IC_{50}) of 4.95 ± 0.33 nM and a selective index (SI) of 456 ± 31. Furthermore, the BoScFv-PE38 exerted a cytotoxic effect through the induction of ATP and ammonia, leading to apoptosis of BoHV-1-infected cells and the inhibition of BoHV-1 replication in MDBK cells. Collectively, we show that BoScFv-PE38 can potentially be employed as a therapeutic agent for the treatment of BoHV-1 infection.

Keywords: *Bovine herpesvirus-1*, immunotoxin, antiviral function, specific binding, targeted cytotoxic effect, apoptosis

INTRODUCTION

Bovine herpesvirus-1 (BoHV-1) belongs to the *Herpesviridae* family in the *Alphaherpesvirinae* subfamily (Muylkens et al., 2007) and is an economically important pathogen that causes infectious bovine rhinotracheitis (IBR) in cattle (Rola et al., 2017; Thakur et al., 2017). BoHV-1 infected animals experience a range of mild to severe clinical syndromes, including rhinotracheitis, vaginitis, balanoposthitis, abortion, conjunctivitis, and enteritis, together with reduced milk production, and weight gain (Raaperi et al., 2014). BoHV-1 pathobiology is somewhat similar to the human herpesvirus 1 (HHV-1), having a short replication cycle and the ability to cause

life-long infection (Levings and Roth, 2013; Zhu et al., 2017). BoHV-1 can also serve as disease model for improving control strategies against *herpesviruses* infecting both humans and animals. Although BoHV-1 vaccines are effective at reducing the clinical impact of BoHV-1 infection, the available vaccines provide suboptimal protection against BoHV-1 in cattle (Muylkens et al., 2007). Therefore, it is necessary to develop antiviral agents that target infected cells to clear virus in host, especially act as a reservoir for spreading virus throughout a herd (Frizzo da Silva et al., 2013). Treatment of viral infections with currently available synthetic drugs possess several deficiencies including toxicity and resistance (Spiess et al., 2016; Khandelwal et al., 2017; Wambaugh et al., 2017), therefore, there is urgency for new and improved antivirals. Recently, immunotoxins against a variety of viruses have been developed, including single-stranded RNA viruses infecting humans, such as HIV, PCV, rabies virus, and herpesvirus, HCMV, EBV and HSV-2 (Mareeva et al., 2010; Chatterjee et al., 2012; Spiess et al., 2017). Immunotoxins, that are chimeric proteins consisting of the antigen-binding fragment (Fab) of an antibody conjugated to a toxin molecule, have shown promise in targeted delivery of antiviral toxins to virus infected cells (Margolis et al., 2016; Spiess et al., 2016). There is growing interest in developing immunotoxins for use in cancer treatment, and lately, the development of a variety of immunotoxins has been reported with the ability to inhibit virus replication and dissemination along with destruction and clearance of infected cells (Mazor et al., 2012; Denton et al., 2014; Chandramohan et al., 2017; Lim et al., 2017; Polito et al., 2017). The major beneficial effect of antibody-conjugated immunotoxins is that they are selective and provide targeted delivery of toxins with minimal side effects to the host (Cai and Berger, 2011; Hou et al., 2016; Müller et al., 2017). Therefore, the target molecule is the major element within the immunotoxin and plays a vital role in targeting virus-infected cells.

The targeting of cell surface antigens or pathogens is usually achieved through the use of their specific monoclonal antibodies (mAbs). The Fab portion of mAbs can be genetically engineered as a recombinant single-/double-chain antibody fragment, or constructed as a single-chain antibody fragment (scFv) for use a as a targeting molecule. These scFv molecules have been used in various immunotoxins due to its high specificity and binding ability. Furthermore, scFv displays good biocompatibility with low antigenicity and may not elicit an immune response when administered to animals and humans (Schotte et al., 2014; Della Cristina et al., 2015; Hanke et al., 2016; Liu B. et al., 2016). Bacterial toxins (*Pseudomonas* exotoxin or *diphtheria* toxin) are most commonly used to prepare immunotoxins, due to irreversibly inhibit protein synthesis in eukaryotic cells via ADP-ribosylation of translation elongation factor 2 (eEF2) (Chatterjee et al., 2012; Spiess et al., 2016).

In our previous study, we demonstrated that scFv targeting of viral glycoprotein D (gD) inhibited the infectivity of BoHV-1 in Madin-Darby bovine kidney (MDBK) cells (Xu et al., 2017). In the present study, we developed BoHV-1-specific scFv that acted as the targeting molecule. Recombinant bacterial toxin derived from *Pseudomonas* exotoxin A (PE38) linked with BoHV-1-specific scFv (BoScFv-PE38) showed immunotoxin activity by binding to BoHV-1 particles in virus-infected cells and exerting a specific cytotoxic effect through induction of high levels of ATP and ammonia production, leading to apoptosis of BoHV-1-infected cells. As a result, replication of BoHV-1 was significantly reduced in MDBK cells.

MATERIALS AND METHODS

Cells and Viruses

MDBK cells and human embryonic kidney HEK293T (293T) cells were purchased from the American Type Culture Collection (Manassas, VA, United States). The MDBK and 293T cell lines were cultured at 37°C in a 5% CO_2 incubator in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum. *Bovine herpesvirus* 1 (BoHV-1) (BK1952) was obtained from the China Veterinary Culture Collection Center (CVCC), Beijing, China, and grown in MDBK cells.

Plasmids and Antibodies

The pET28a expression system was obtained from GE Healthcare; the pEGFP-N1 vector was obtained from Clontech; The DNA encoding for segment from 259 amino acids to 345 amino acids of glycoprotein D (AFB76672.1) was amplified by polymerase chain reaction with the primers as following: Sense primer: 5′-GAATTCATGGAGGAGTCGAAGGGC-3′ and anti-sense primer:5′-CTCGAGGATGGCTTCGAGGCTCG-3′, and the DNA fragment was cloning into the pEGFP-N1 vector for construction of the pEGFP-N1-gD, which was used to efficiently express green fluorescent protein (GFP) fused gD protein in 293T cell. Calf antiserum against BoHV-1 was from China Veterinary Culture Collection Center. A mouse anti-His monoclonal antibody (McAb), Alexa Fluor 555- conjugated anti-His antibody, FITC-labeled goat anti-mouse antibody and TRITC-labeled goat anti-mouse antibody were purchased from ThermoFisher Scientific (United States), and a FITC-labeled rabbit anti-bovine antibody was purchased from BioVision (United States). The antibodies against PARP-1, Bcl-2, Bid, caspase-3, caspase-8, caspase-9 and β-actin were purchased from ABclonal Biotech (China). The BoScFv-PE38 was labeled with horseradish peroxidase (HRP) by Sangon Biotech Co., Ltd. (China).

Sequence Analysis and Expression of the Immunotoxin

The protein sequence of *Pseudomonas* exotoxin A (PE38) was downloaded from the NCBI database[1]; sequences of a truncated version of PE38, the BoHV-1 ScFv protein and the linker peptide are obtained according to our earlier study (Xu et al., 2017), which were listed in Supplementary Table S1. The domains of BoScFv-PE38 were analyzed using the PROSITE

[1]https://www.ncbi.nlm.nih.gov/protein/553773623

database[2]. The full-length nucleotide sequence of BoScFv-PE38 was optimized and synthesized by Shanghai Sangon Biotech Co., Ltd. (Shanghai, China). The fusion gene BoScFv-PE38 was cloned into expression vector pET28a and expressed in *Escherichia coli* BL21 (DE3) (Novagen, EMD Chemicals, Inc., Madison, WI, United States). The recombinant His tagged BoScFv-PE38 was purified via nickel affinity chromatography as described by Della Cristina et al. (2015). The endotoxin was removed from purified BoScFv-PE38 using the Detoxi-Gel^{TM} Endotoxin Removing Columns Kit (ThermoFisher Scientific, United States), and the endotoxin residue in purified BoScFv-PE38 was detected with the ToxinSensor^{TM} Chromogenic LAL Endotoxin Assay Kit (Kingsy Biotechnology, Nanjing, China). The purified BoScFv-PE38 (endotoxin < 0.0068 EU/ml) was dissolved in phosphate-buffered saline (PBS, pH 7.4) solution and stored at −20°C.

Measurement of the Dissociation Constants of BoScFv-PE38

The Kd of BoScFv-PE38 was measured via ELISA. Briefly, 96-well microplates were coated with the BoHV-1 or gD protein (expressed and purified from *E. coli*) and blocked with 3% bovine serum albumin (BSA). Serial dilutions of BoScFv-PE38 labeled with horseradish peroxidase (HRP) (HRP-BoScFv-PE38) (0–100 nM for the gD protein, 0–1500 nM for BoHV-1) were added to the wells, followed by incubation at 37°C for 60 min. Next, the tetramethylbenzidine (TMB) (Sigma) substrate was added, followed by incubation for 10 min. Then, stop buffer (2 M sulfuric acid) was added to stop the reaction. Finally, the optical densities were read at 450 nm, and the equation Y = Bmax X/(Kd+X) was used to obtain the saturation curve and Kd of BoScFv-PE38, employing GraphPad Prism 5.0. Y represents the mean OD450 nm value; Bmax is the maximal OD450 nm value; and X is the concentration of BoScFv-PE38.

Immunofluorescence Assay and Confocal Laser Scanning Microscopy

The Immunofluorescence assay (IFA) was performed as described previously (Keuser et al., 2004). Briefly, 293T cells were seeded onto cover slips in six-well plates and cultured to 70% confluency at 37°C over 18–24 h. Then, the PEGF-N1 and PEGF-N1-gD plasmids were transfected with Lipofectamine 3000 (Life Technology, United States) according to the manufacturer's instructions. After 6 h, BoScFv-PE38 (1 μM) was added to the culture medium, and the 293T cells were cultured for an additional 24 h. Then, the 293T cells were fixed with 4% paraformaldehyde, blocked with 3% bovine serum albumin, and incubated with the Alexa Fluor 555-anti-His antibody at 37°C for 1 h. Finally, the nuclei were counterstained with DAPI (blue), and the cell samples were examined under a fluorescence microscope (Leica EL 6000). The co-localization of gD with BoScFv-PE38 or the karyomorphism in 293T cells was observed under a confocal laser scanning microscopy (CLSM, Leica).

MDBK cells were seeded on cover slips in six-well plates and cultured to 50% confluency at 37°C for 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1). After 1.5 h, the culture medium was replaced with fresh culture medium containing 1% FBS and BoScFv-PE38 (1 μM), and the MDBK cells were cultured for an additional 24 h. Next, the MDBK cells were fixed with 4% paraformaldehyde and blocked with 3% bovine serum albumin. To detect the binding of BoScFv-PE38 in BoHV-1-infected cells, the cells were incubated with the anti-His McAb at 37°C for 1 h, followed by incubation with the FITC-labeled goat anti-mouse antibody at 37°C for 1 h. Sequently, in order to monitor the localization of BoHV-1 or BoHV-1 gD with BoScFv-PE38, the cells were incubated with the anti-gD McAb or anti-BoHV-1 Bovine serum at 37°C for 1 h. After washing three times with PBS, the cells were incubated with the FITC-labeled goat anti-mouse antibody or FITC-labeled rabbit anti-bovine antibody, respectively, at 37°C for 1 h. The cells were next washed three times with PBS and subsequently incubated with Alexa Fluor 555-conjugated anti-His antibody at 37°C for 1 h. Finally, the nuclei were counterstained with DAPI (blue), and cell samples were examined with a fluorescence microscope (Leica EL 6000). The co-localization of gD and BoScFv-PE38 and the karyomorphism of MDBK cells were observed under a CLSM (Leica).

Cytotoxicity of BoScFv-PE38 to 293T Cells Expressing gD

293T cells were seeded in 96-well plates and cultured to 70% confluency at 37°C for 18–24 h. Then, the PEGF-N1 and PEGF-N1-gD plasmids were transfected with Lipofectamine 3000 (Life Technology, United States) according to the manufacturer's instructions. After 6 h, BoScFv-PE38 (0–1.0 μM) was added to the culture medium, followed by cultivation at 37°C for 24 h. Cell proliferation was tested with the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (MTS) Kit (Promega) according to the manufacturer's instructions, and the OD490 values of the test wells were read with a microplate reader, the cellular viability was calculated according to the equation: Cellular viability (%) = OD490 value$_{(\text{Treatment group})}$/OD490 value$_{(\text{Control group})}$ × 100. All tests were performed in triplicate.

Cytotoxicity of BoScFv-PE38 to MDBK Cells Infected With BoHV-1

To determine a proper dose of BoHV-1 for inoculating MDBK cells, the cells were seeded in 96-well plates and cultured to 90% confluency. Then, the MDBK cells were inoculated with a 0.01–10 multiplicity of infection (MOI) of BoHV-1 at 37°C for 1.5 h. Then, the culture medium was replaced with fresh culture medium containing 1% FBS and BoScFv-PE38 (1.00 μM), and the cells were cultured for 24 h at 37°C. Cellular viability in each plate was calculated via the MTS assay in quadruplicate. Finally, the specific cytotoxicity of BoScFv-PE38 was also evaluated using the MTS assay as described above with 1 MOI of BoHV-1 and 0–1 μM BoScFv-PE38. Each experiment was repeated four times.

Analysis of Apoptosis

Titration of ATP and ammonia in MDBK cells was performed as follows. MDBK cells were seeded on cover slips in 96-well

[2]http://www.expasy.ch/tools/scanprosite/

plates and cultured to 90% confluency at 37°C for 18–24 h. The MDBK cells were then infected with BoHV-1 (MOI = 1) and cultivated at 37°C for a further 1.5 h. Then, the culture medium was replaced with fresh culture medium containing 1% FBS and BoScFv-PE38 (0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.50, 1.00, and 2.00 μM), and the cells were cultured for an additional 24 h. The concentrations of ATP and ammonia were measured with an ATP determination kit (Sigma) and an ammonia assay kit (Sigma) according to the manufacturer's instructions.

The TUNEL assay was employed as described previously (Hohensinner et al., 2017). MDBK cells were seeded in 24-well plates and cultured to 90% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1) and cultivated at 37°C for 1.5 h, followed by culture in medium containing 1% FBS and BoScFv-PE38 (0.125, 0.25, 0.5, and 1 μM) for 24 h. Apoptosis-positive MDBK cells were detected with the *In Situ* Cell Death Detection Kit, AP (Roche) according to the manufacturer's instructions, and the apoptotic rates of MDBK cells in the presence of different concentration of BoScFv-PE38 were determined.

Western Blot Analysis

The purified BoHV-1 gD protein (20 μg) or BoHV-1 (20 μg) was separated through sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore Schwalbach, Germany). Next, the membrane was blocked with 5% bovine skimmed milk and then incubated with BoScFv-PE38 labeled with horseradish peroxidase (HRP) at 37°C for 1.5 h. Thereafter, the membrane was washed three times with PBS containing 0.5% Tween-20. The membrane was finally developed with ECL solution (ThermoFisher Scientific, United States) and visualized with the Odyssey Infrared Imaging System (LI-COR Biosciences).

The expression of apoptotic proteins in BoHV-1-infected MDBK cells after treatment with BoScFv-PE38 was analyzed as described by Decker et al. (2004). Fifty micrograms of protein was separated via 12% SDS-PAGE and transferred to PVDF membranes. The samples were subsequently hybridized with antibodies against PARP-1, Bcl-2, Bid, caspase-3, caspase-8, caspase-9 and β-actin (ABclonal Biotech Co., Ltd, China). Blots were developed using Super Signal chemoluminescent substrates (Pierce, KMF GmbH, St. Augustin, Germany).

Plaque Reduction Assay

The antiviral activity of BoScFv-PE38 was further evaluated through the plaque reduction test (PRT) as previously described (Levings et al., 2015). BoScFv-PE38 (0.25, 0.5, 1, and 2 μM) was mixed with 10–50% tissue culture infective dose ($TCID_{50}$) of BoHV-1 in an equal volume. After incubation at 37°C under 5% CO_2 for 1 h, 1 ml of the mixture was inoculated in triplicate into the wells of a six-well plate containing a confluent monolayer of MDBK cells. The plates were subsequently incubated at 37°C under 5% CO_2 for 1.5 h with intermittent rocking. Then, an agarose overlay was added to the infected cell monolayer. After the agarose became solid, the plates were placed upside down and were further incubated for 48 h. When viral plaques became visible, the cells were fixed with 4% formaldehyde and stained with 0.1% toluidine blue in saline, followed by visual counting of viral plaques.

Infectious Center Assay

The infectious center assay (ICA) was employed as reported previously (Geoghegan et al., 2015). MDBK cells were handled, as described above, and treated with BoScFv-PE38 (0.125, 0.25, 0.5, and 1 μM) for 24 h. All cultures were harvested and subjected to repeated freeze-thaw three times at 2, 10, 16, and 24 h. The culture mixture (harvested at 2 h) was diluted to 1:100–1:1000. The diluted culture mixtures were used to infect MDBK cells in 6-well plates for 1.5 h. An agarose overlay was then added to the infected cell monolayer, and after the agarose solidified, the plates were placed upside down and further incubated for 48 h, and viral plaques were counted visually.

50% Cytotoxic Concentration, 50% Inhibitory Concentration and Selective Index

To evaluate the cytotoxic concentration of BoScFv-PE38, MDBK cells were seeded in 96-well plates and cultured to 90% confluency at 37°C for 18–24 h. Then, BoScFv-PE38 (0.1263–20.0 μM) was added to the culture medium, followed by cultivation at 37°C for 24 h. Cell proliferation was tested with the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (MTS) Kit (Promega) according to the manufacturer's instructions, and the OD490 values of the test wells were read with a microplate reader for calculation of the cellular viability, and the cytotoxic concentration and 50% cytotoxic concentration (CC_{50}) was calculated. To determine the 50% inhibitory concentration (IC_{50}) of BoScFv-PE38, twofold serial dilutions of BoScFv-PE38 (initial 250 nM) were incubated with BoHV-1 (MOI = 1) for 1 h at 37°C. Thereafter, the mixtures were added to MDBK cells, followed by culture for 72–96 h, and the cytopathic effect was observed and calculated as the IC_{50} value. The selective index (SI) was calculated according to the equation: $SI = CC_{50}/IC_{50}$. All tests were performed in triplicate (Supplementary Table S2).

Statistics

All the statistical analyses were performed via analysis of variance (ANOVA) using SPSS software, version 18.0, and the fitting curves were drawn with GraphPad Prism 5.0. *P*-values < 0.01 were considered statistically significant. $P > 0.05$ represents no statistically significant differences. $P < 0.05$ represents statistically significant differences. $P < 0.01$ represents statistically significant differences. $P < 0.001$ represents statistically significant differences.

RESULTS

Construction of BoScFv-PE38 Immunotoxin

A DNA fragment containing the open reading frame of PE38 was synthesized according to the published sequence of *Pseudomonas*

exotoxin A (PE38), derived from *Pseudomonas aeruginosa*, without an intrinsic cell-binding domain (**Figure 1A**). The modified PE38 nucleotide sequence was then fused with a gene encoding a scFv mAb that has high specificity toward the gD protein of BoHV-1; the result was a 1863 bp construct of BoScFv-PE38 immunotoxin (**Figure 1B** and Supplementary Table S1). In the BoScFv-PE38 construct, the Exotox-A-binding region of PE38 was replaced with an scFv against the gD protein of BoHV-1, but the Exotox-A-targeting and Exotox-A catalytic domains were maintained and ligated with the binding domain of scFv via a 3 (G_4S) linker (**Figure 1C**).

BoScFv-PE38 Immunotoxin Showed Specific Binding Affinity for BoHV-1

To assess the binding affinity of BoScFv-PE38 immunotoxin to the BoHV-1, the DNA fragment of BoScFv-PE38 was cloned into the pET28a plasmid and expressed in *Escherichia coli* BL21(ED3) (**Figure 2A**). The BoScFv-PE38 immunotoxin was then purified via affinity chromatography (**Figure 2B**). The western blot analysis showed that both the recombinant gD protein produced in *E. coli*. and the wild type gD protein derived from BoHV-1 virus had specific binding affinity for purified BoScFv-PE38 immunotoxin (**Figures 2C,D**). Enzyme-linked immunosorbent assay (ELISA) analysis also showed very high binding avidity of BoScFv-PE38 immunotoxin with the recombinant gD protein and BoHV-1 (**Figures 2C,D**). The dissociation constants (Kds) of BoScFv-PE38 binding to gD and BoHV-1 were 12.81 ± 2.24 and 97.63 ± 10.88, respectively (**Figures 2E,F**), indicating that the BoScFv-PE38 protein had specific and high binding avidity for both the BoHV-1 gD and BoHV-1 virus particles. To further test the ability of BoScFv-PE38 immunotoxin to capture BoHV-1 virus, a double sandwich ELISA was established by coating a 96-well microplate with different doses of BoScFv-PE38. Then, the amount of BoHV-1 antigen that was captured was titrated via ELISA. As shown in **Figure 2G**, there was a positive correlation between the amounts of BoHV-1 antigen captured and the concentrations of BoScFv-PE38 immunotoxin. The comparative control assays showed no significant differences among the tested dilutions (1:300–1:2400) of standard bovine serum against BoHV-1 (**Figure 2G**).

BoScFv-PE38 Immunotoxin Interacts With BoHV-1 gD Protein Within Cells

To determine whether BoScFv-PE38 immunotoxin could specifically target gD protein of BoHV-1 within the cellular organelles, the pEGF-N1-gD construct was engineered and transfected into 293T cells (Supplementary Figure S1). Then, 293T cells expressing the BoHV-1 gD protein were detected by staining with BoScFv-PE38 and an Alexa Fluor 555-conjugated mouse anti-His antibody. Indirect immunofluorescence assays (IFAs) and confocal microscopic assays confirmed that both the gD protein (labeled for red fluorescence) and the BoScFv-PE38 immunotoxin (labeled for green fluorescent protein (GFP)) were co-localized inside the cells (**Figures 3A,B**).

BoScFv-PE38 Immunotoxin Has Specific Cytotoxic Activity for Cells Expressing BoHV-1 gD Protein

The specific cytotoxic activity of BoScFv-PE38 immunotoxin for cells expressing BoHV-1 gD protein was evaluated using a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxyme-thoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay. The protein BoScFv-PE38 was added to 293T cells expressing either GFP-fused gD or GFP, followed by co-culture for 18 h. The results showed that BoScFv-PE38 could significantly inhibit the proliferation of 293T cells expressing BoHV-1 gD protein in a dose-dependent manner, but there was no

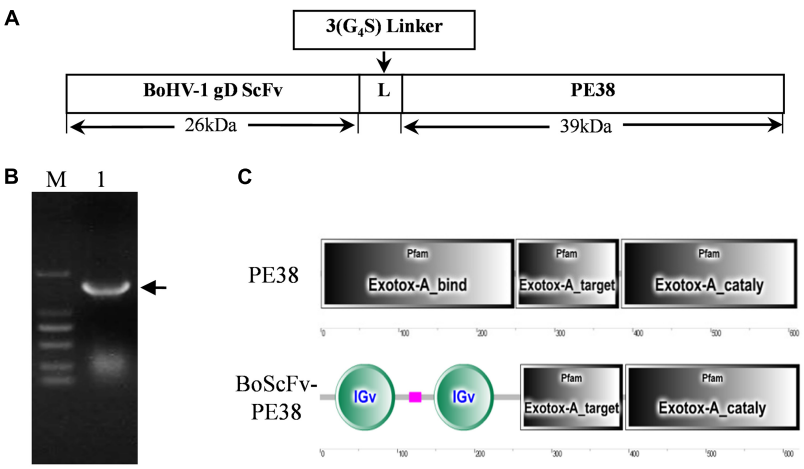


FIGURE 1 | Nucleotide sequence structure of BoScFv-PE38. **(A)** Schematic diagram of the recombinant immunotoxin BoScFv-PE38. BoHV-1 gD scFv, BoHV-1-specific nanobody; 3(G_4S) linker, flexible linkers consisting of glycine and serine residues; PE38, pseudomonas exotoxin A (PE38) protein without the Exotox-A-binding region. **(B)** The BoScFv-PE38 nucleotide sequence. M represents the DNA molecular weight marker; 1 represents the BoScFv-PE38 nucleotide sequence; and the arrow indicates the size of the BoScFv-PE38 nucleotide sequence. **(C)** Structural sequence analysis of BoScFv-PE38 and PE38.

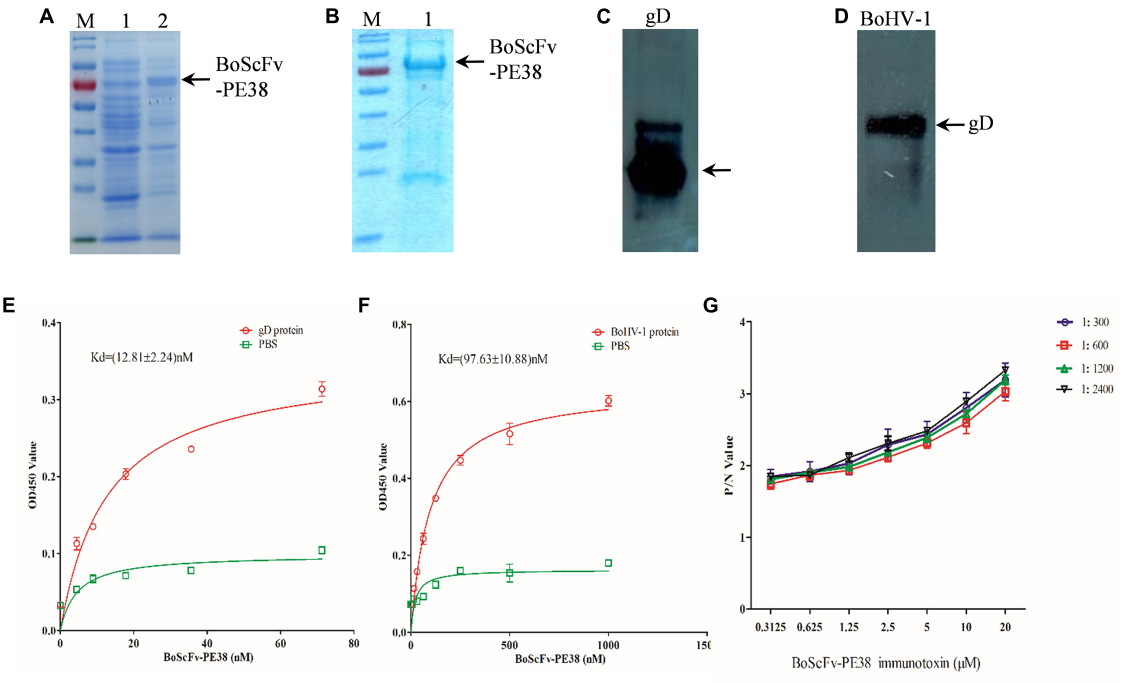


FIGURE 2 | Expression and binding analysis of BoScFv-PE38. **(A)** Expression of BoScFv-PE38 in *E. coli*. M: PageRuler Prestained Protein Ladder; 1: the pET28a vector was transformed into *E. coli* cells, and expression was induced with IPTG; 2: the pET28a- BoScFv-PE38 vector was transformed into *E. coli* cells, and expression was induced with IPTG. The arrow represents the recombinant ScFv. **(B)** Purification of BoScFv-PE38 from *E. coli* via nickel affinity chromatography. M: PageRuler Prestained Protein Ladder. 1 indicates purified BoScFv-PE38, and the arrow represents purified ScFv. **(C)** The reaction between gD and BoScFv-PE38 was detected by western blotting. Recombinant gD protein detected by BoScFv-PE38-HRP. **(D)** The reaction between BoHV-1 and BoScFv-PE38 was detected by western blotting. BoHV-1: BoHV-1 was detected by BoScFv-PE38-HRP. **(E)** Measurement of the Kd of BoScFv-PE38 binding to the gD protein. First, 96-well plates were coated with gD (1 μg). The plates were then incubated with BoScFv-PE38-HRP and developed with TMB substrate solution. The reaction was stopped with 2 M sulfuric acid, and the optical density was read at 450 nm. The equation Y = BmaxX/(Kd + X) was used to obtain the saturation curve and the Kd of the ScFv-gD interaction using GraphPad Prism 5.0. Y represents the mean OD450 value; Bmax is the maximal OD450 value; and X is the concentration of BoScFv-PE38. **(F)** Measurement of the Kd of BoScFv-PE38 binding to BoHV-1. Ninety-six-well plates were coated with BoHV-1 (1 μg), and the Kd of the BoScFv-PE38-BoHV-1 interaction was measured as described above. **(G)** Binding affinity of BoScFvScFv-PE38 was detected using an antigen capture ELISA. Ninety-six-well microplates were coated with different concentrations of BoScFvScFv-PE38 and blocked with 3% BSA. The BoHV-1 antigen (standard serially diluted BoHV-1-positive serum; 1:300–1:2400), and HRP-conjugated goat anti-bovine IgG were added, and the plates were developed with the TMB substrate (Sigma). The reaction was stopped with 2 M sulfuric acid, and the optical density was read at 450 nm. The P/N values were analyzed using GraphPad Prism 5.0.

obvious cytotoxic effect on control GFP-expressing 293T cells (**Figure 3C**). These results were supported by the observation of cell karyomorphism, in that there was an abnormal morphology of BoScFv-PE38-treated gD-expressing 293T cells compared with PBS-treated gD-expressing 293T cells (**Figure 3D**).

BoScFv-PE38 Immunotoxin Captures BoHV-1 in MDBK Cells

The ability of BoScFv-PE38 immunotoxin to specifically recognize and interact with BoHV-1 virus particles within infected cells was investigated through IFA. The results showed that BoScFv-PE38 immunotoxin selectively internalized into the MDBK cells infected with BoHV-1 (green fluorescence) rather than negative control cells (**Figure 4A**), indicating that BoScFv-PE38 immunotoxin had efficient capability to target and internalize BoHV-1 in MDBK cells. To verify this, we observed the distribution of BoHV-1 virus particles bound to BoScFv-PE38 immunotoxin or BoHV-1 virus-specific polyclonal antibodies within an individual cell using a confocal laser-scanning microscope (CLSM). The results indicated that BoHV-1 virions recognized by BoScFv-PE38 immunotoxin were localized in the cytoplasm and to the periphery of the nucleus. In contrast, polyclonal antibodies were only able to detect virus presence at the cell surface (**Figure 4B**). These analyses also revealed that the BoScFv-PE38 immunotoxin was widely distributed on the cell surface as well as in cytoplasm of BoHV-1 infected MBCK cells (**Figure 4C**). Taken together, these results suggest that BoScFv-PE38 immunotoxin is readily internalized into the BoHV-1 infected cells following binding to the BoHV-1 gD protein expressed on the cell surface.

BoScFv-PE38 Immunotoxin Displays Specific Cytotoxic Effect for BoHV-1-Infected MDBK Cells

Since BoScFv-PE38 had the ability to bind to MDBK cells infected with BoHV-1, it was important to determine whether

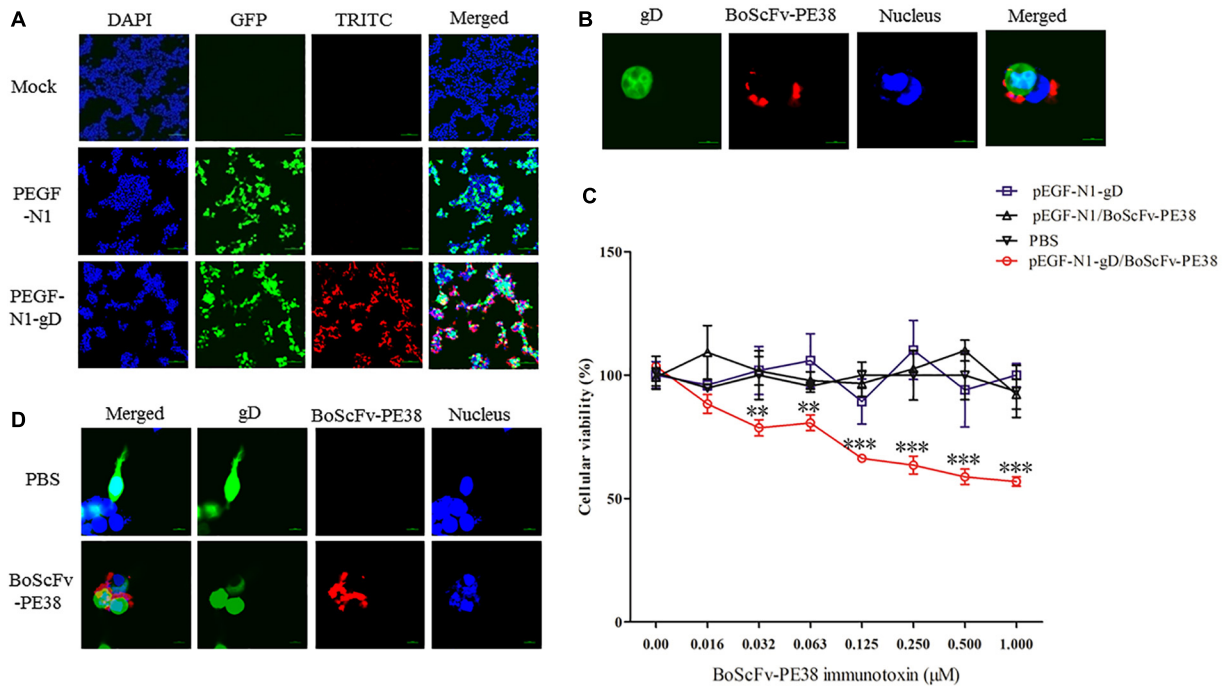


FIGURE 3 | Binding specificity and cytotoxicity of BoScFv-PE38 in gD-expressing 293T cells. **(A)** Binding specificity of BoScFv-PE38 in gD-expressing 293T cells determined via IFA. 293T cells were seeded on cover slips in six-well plates and cultured to 70% confluency at 37°C over 18–24 h. Then, the PEGF-N1 and PEGF-N1-gD plasmids were transfected into the cells. After 6 h, BoScFv-PE38 (1 μM) was added to the culture medium, and the 293T cells were cultured for an additional 24 h. The 293T cells were subsequently fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, blocked with 3% BSA, and incubated with the Alexa Fluor 555-anti-His antibody at 37°C for 1 h. Finally, the nuclei were counterstained with DAPI (blue), and the cell samples were examined with a fluorescence microscope (Leica EL 6000). **(B)** Localization of BoScFv-PE38 in gD-expressing 293T cells determined via CLSM. The 293T cells were treated as described above. Then, 293T cell samples were examined with a confocal laser-scanning microscope (Leica). **(C)** Cytotoxicity of BoScFv-PE38 in gD-expressing 293T cells. 293T cells were seeded on cover slips in 96-well plates and cultured to 70% confluency at 37°C over 18–24 h. Then, the PEGF-N1 and PEGF-N1-gD plasmids were transfected with Lipofectamine 3000 (Life Technology, United States) according to the manufacturer's instructions. After 6 h, BoScFv-PE38 (0–1.0 μM) was added to the culture medium, followed by cultivation at 37°C for 24 h. Then, cell proliferation activity was measured with the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (MTS) Kit (Promega) according to the manufacturer's instructions, and the OD490 values of the test wells were read with a microplate reader for calculation of the cellular viability. All the statistics were analyzed through analysis of variance (ANOVA) using SPSS software, version 18.0, and the fitting curves of various concentrations of BoScFv-PE38 were drawn with GraphPad Prism 5.0, respectively. "**" represents statistically significant differences ($P < 0.01$). "***" represents statistically very significant differences ($P < 0.001$). **(D)** Karyomorphism of gD-expressing 293T cells after treatment with BoScFv-PE38. 293T cells were seeded on cover slips in six-well plates and cultured to 70% confluency at 37°C over 18–24 h. Then, the PEGF-N1-gD plasmid was transfected into the cells. After 6 h, BoScFv-PE38 (1 μM) was added to the culture medium (PBS treated, as the negative control), and the 293T cells were cultured for 24 h. The 293T cells were then fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, blocked with 3% BSA, and incubated with the Alexa Fluor 555-anti-His antibody at 37°C for 1 h. The nuclei were counterstained with DAPI (blue). Then, the karyomorphism of 293T cell samples was examined with a confocal laser-scanning microscope (Leica).

BoScFv-PE38 immunotoxin had selective cytotoxic effect on cells infected with BoHV-1. Initially, we estimated the cytotoxic concentration of BoScFv-PE38 immunotoxin required to induce specific cytotoxic effect on BoHV-1 infected MDBK cells. The MTS assays showed a reduced cell viability when BoScFv-PE38 immunotoxin concentration was increased to 1.25 μM in culture medium, indicating that the cytotoxic concentration of BoScFv-PE38 immunotoxin for MDBK cells should not exceed to 1.25 μM and the cytotoxic concentration of 50% (CC_{50}) was 2.25 μM (**Figure 5A**). The cell proliferation assays clearly demonstrated that a dose of 1.00 μM BoScFv-PE38 immunotoxin can significantly inhibit ($P < 0.001$) the proliferation the MDBK cells infected with range of 0.05–2.5 multiplicities of infection (MOIs) of BoHV-1(**Figure 5B**). Based on these results, we infected MDBK cells with 1.0 MOI of BoHV-1 and then treated the cells with gradient concentrations of BoScFv-PE38 immunotoxin. The data revealed that the proliferation of MDBK cells was continually reduced with the increase of BoScFv-PE38 immunotoxin concentration (**Figure 5C**). Suggesting that the BoScFv-PE38 immunotoxin exerted dose-dependent cytotoxic effects on BoHV-1-infected MDBK cells. To confirm the above findings, we observed the karyomorphism of MDBK cells infected with BoHV-1 after being treated with BoScFv-PE38 immunotoxin. Confocal microscopy images corroborated above findings indicating the nuclei of BoHV-1-infected MDBK cells were defective after treatment with 1.00 μM BoScFv-PE38, whereas untreated cells maintained normal morphology until 18 h post-infection despite infection with equal infectious dose of BoHV-1 (**Figure 5D**).

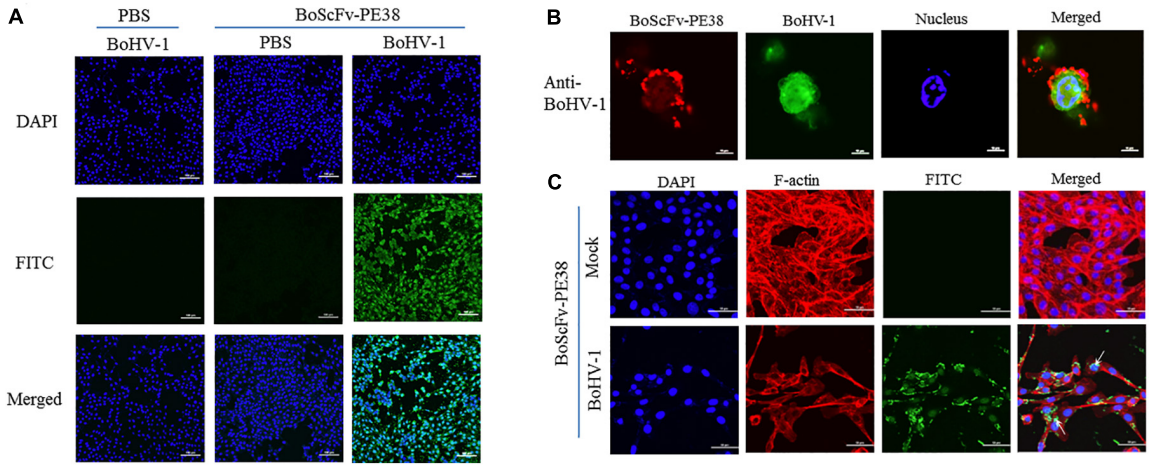


FIGURE 4 | Binding specificity and internalization of BoScFv-PE38 in BoHV-1-infected MDBK cells. **(A)** Recognition of BoScFv-PE38 to BoHV-1 infected in MDBK cells was determined via IFA. MDBK cells were seeded on cover slips in six-well plates and cultured to 50% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1). After 1.5 h, the culture medium was replaced by new culture medium containing 1% FBS and BoScFv-PE38 (1 μM), and the MDBK cells were cultured for 24 h. The MDBK cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, blocked with 3% bovine serum albumin, incubated with the anti-His McAb for 1 h at 37°C, and then incubated with the FITC-goat anti-mouse antibody for 1 h at 37°C. The nuclei were counterstained with DAPI (blue). The cell samples were examined with a fluorescence microscope (Leica EL 6000). **(B)** Location of BoScFv-PE38 in BoHV-1-infected MDBK cells was determined via CLSM. MDBK cells were seeded on cover slips in six-well plates and cultured to 50% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1). After 1.5 h, the culture medium was replaced with new culture medium containing 1% FBS and BoScFv-PE38 (1 μM), and the MDBK cells were cultured for a further 24 h. The MDBK cells were then fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, blocked with 3% bovine serum albumin, incubated with the anti-gD mAb or anti-BoHV-1 bovine serum for 1 h at 37°C, washed three times with PBS, incubated with the FITC-labeled goat anti-mouse antibody or the FITC-labeled rabbit anti-bovine antibody at 37°C for 1 h, washed three times with PBS once more, and finally incubated with the Alexa Fluor 555- anti-His antibody at 37°C for 1 h. The nuclei were counterstained with DAPI (blue). The cell samples were observed under a confocal laser-scanning microscope (Leica). **(C)** The internalization of BoScFv-PE38 entrapped by BoHV-1. MDBK cells were infected with BoHV-1 (MOI = 1). After 1.5 h, the culture medium was replaced with new culture medium containing 1% FBS and BoScFv-PE38 (1 μM), and the MDBK cells were cultured for 24 h. MDBK cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, and blocked with 3% bovine serum albumin. The cells were then incubated with the anti-His mAb for 1 h at 37°C, washed three times with PBS, and incubated with the FITC-labeled goat anti-mouse antibody at 37°C for 1 h. Cellular F-actin was stained with Alexa Fluor 555-conjugated phalloidin. The nuclei were counterstained with DAPI (blue). Normal MDBK cells were used as the negative control. Cell samples were examined with a confocal laser-scanning microscope (Leica). The white arrows indicate BoScFv-PE38 in the nucleus and cytoplasm of BoHV-1-infected MDBK cells.

BoScFv-PE38 Immunotoxin Induces Apoptosis of BoHV-1-Infected Cells

To explore the mechanism of the specific cytotoxic effect of BoScFv-PE38 immunotoxin on the BoHV-1 infected MDBK cells, we analyzed the concentrations of adenosine triphosphate (ATP) and ammonia, which are the major indicators associated with cytotoxicity, in BoScFv-PE38 treated, BoHV-1-infected MDBK cells (Leist et al., 1997; Cheng et al., 2015). The results showed that the concentrations of ATP and ammonia in BoHV-1-infected MDBK cells treated with BoScFv-PE38 were significantly higher than in cells that were infected with BoHV-1 alone, non-infected cells treated with BoScFv-PE38, and negative cells (mock treated with PBS). Additionally, the titers of ATP and ammonia increased with the concentration of BoScFv-PE38 within 0.015625–1 μM or 0.015625–2.00 μM, respectively (**Figures 6A,B**). This finding demonstrated that BoScFv-PE38 immunotoxin increased the production of ATP and ammonia in BoHV-1-infected MDBK cells. Observed high levels of ATP or ammonia imply that cell death was induced by apoptosis. The apoptotic cells were further subjected to terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assays. Increased amount of apoptotic bodies were observed in BoHV-1-infected and BoScFv-PE38 immunotoxin treated MDBK cells compared with mock treated BoHV-1 infected cells (**Figure 6C**). The rate of apoptosis induced by BoScFv-PE38 immunotoxin treated BoHV-1infected cells was significantly higher than in the mock PBS-treated control or following induction by BoScFv-PE38 immunotoxin or BoHV-1 individually. Additionally, there was a dose-dependent increase in the rate of apoptotic cells induced by BoScFv-PE38 immunotoxin bound with BoHV-1 (**Figure 6D**). These results were further supported by the analysis of apoptotic proteins including PRAP-1, Bcl-2, Bid, Caspase 8 and Caspase 3. The results clearly revealed increased cleavage of pro-apoptotic proteins Bid in BoHV-1-infected MDBK cells after treatment with BoScFv-PE38 immunotoxin, while the expression of the anti-apoptosis protein Bcl-2 was relatively decreased (**Figures 6E,F**).

BoScFv-PE38 Effectively Inhibits the Infectivity of BoHV-1 in MDBK Cells

The observed data demonstrated that BoScFv-PE38 immunotoxin specifically targets BoHV-1-infected cells and exerts cytotoxic effects. Therefore, we examined the ability of BoScFv-PE38 immunotoxin to inhibit the infectivity of BoHV-1

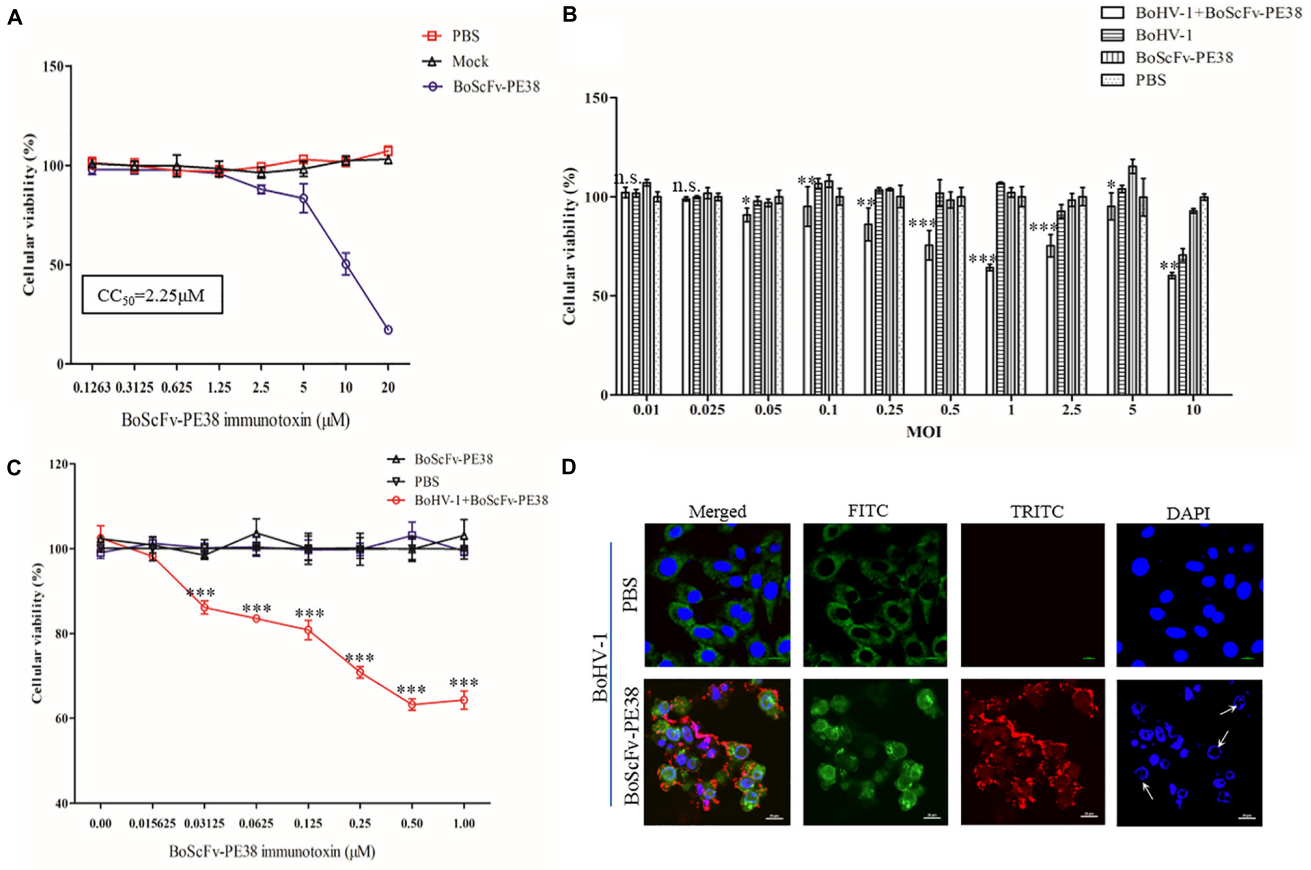


FIGURE 5 | Cytotoxicity of BoScFv-PE38 in BoHV-1-infected MDBK cells. **(A)** Determination of the cytotoxic concentration of BoScFv-PE38 in MDBK cells. MDBK cells were seeded in 96-well plates and cultured to 90% confluency at 37°C for 18–24 h. Then, BoScFv-PE38 (0.1263–20.0 μM) was add to the culture medium, followed by cultivation at 37°C for 24 h, and the cellular viability was measured. **(B)** Cytotoxicity of BoScFv-PE38 in MDBK cells infected with different MOIs of BoHV-1. MDBK cells were seeded in 96-well plates and cultured to 90% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI was from 0.01 to 10), added to the culture medium, and cultivated 37°C for 1.5 h. The culture medium was subsequently replaced with new culture medium containing 1% FBS and BoScFv-PE38 (1 μM); the MDBK cells were cultured for an additional 24 h; and cellular viability was measured. **(C)** Dose-dependent cytotoxicity of BoScFv-PE38 in BoHV-1-infected MDBK cells. MDBK cells were seeded on cover slips in 96-well plates and cultured to 90% confluency at 37°C for 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1), added to the culture medium, and cultivated at 37°C for 1.5 h. The culture medium was subsequently replaced with new culture medium containing 1% FBS and BoScFv-PE38 (1 μM); the MDBK cells were cultured for 24 h; and cellular viability was measured. All statistical analyses were performed via analysis of variance (ANOVA) using SPSS software, version 18.0, and the fitting curves of various concentrations of BoScFv-PE38 were drawn with GraphPad Prism 5.0. "n.s." represents no statistically significant differences ($P > 0.05$). "*" represents statistically significant differences ($P < 0.05$). "**" represents statistically significant differences ($P < 0.01$). "***" represents statistically significant differences ($P < 0.001$). **(D)** Karyomorphism of BoHV-1-infected MDBK cells after treatment with BoScFv-PE38. MDBK cells were seeded on cover slips in six-well plates and cultured to 50% confluency at 37°C for 18–24 h. Then, the MDBK cells were infected with BoHV-1(MOI = 1). After 1.5 h, the culture medium was replaced with new culture medium containing 1% FBS and BoScFv-PE38 (1 μM) or PBS as the negative control, and the MDBK cells were cultured for 24 h. The MDBK cells were subsequently fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 for 1 h, and blocked with 3% bovine serum albumin. The cells were then incubated with anti-BoHV-1 bovine serum for 1 h at 37°C, washed three times with PBS, incubated with the FITC-labeled Rabbit anti-bovine antibody at 37°C for 1 h, washed three times with PBS once more, and finally incubated with the Alexa Fluor 555-anti-His antibody at 37°C for 1 h. Nuclei were counterstained with DAPI (blue). Cell samples were examined with a fluorescence microscope (Leica EL 6000). The karyomorphism of MDBK cells was observed under a CLSM (Leica).

in MDBK cells using micro-neutralization tests and plaque reduction assays. As shown in **Figure 7A**, the BoHV-1 plaque count in MDBK cells was significantly reduced ($P < 0.01$) following pre-incubation of virus with varying concentrations of BoScFv-PE38 immunotoxin for 1 h compared with mock PBS treatment (**Figure 7B**). The inhibitory effect of BoScFv-PE38 immunotoxin on BoHV-1 replication was also evaluated via the infectious center assay (ICA). The results showed that there were a marked decrease in BoHV-1 virus replication in MDBK cell cultures supplemented with a varying of amount (1, 0.5, 0.25, and 0.125 μM) of BoScFv-PE38 immunotoxin (**Figure 7C**). The estimated 50% inhibitory concentration (IC_{50}) of BoScFv-PE38 immunotoxin was 4.95 ± 0.33 nM observed within 24 h of infection using viral plaque reduction assays. These results illustrated that BoScFv-PE38 effectively inhibited the infectivity of BoHV-1 in MDBK cells. From these results, selective index (SI) can be inferred to be 456 ± 31 in terms of the IC_{50} value of the cytotoxic concentration (Supplementary Table S2).

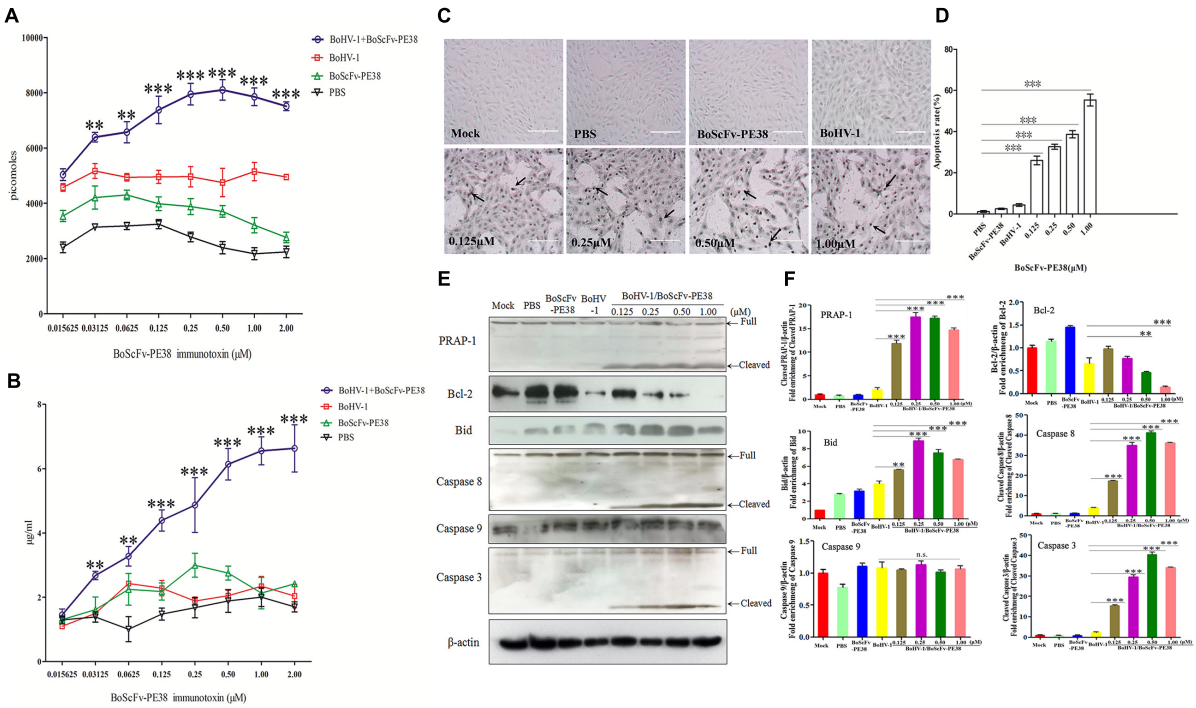


FIGURE 6 | Apoptosis of BoHV-1-infected MDBK cells was induced by BoScFv-PE38. **(A)** The titration of ATP in BoHV-1-infected MDBK cells treated with BoScFv-PE38. MDBK cells were seeded in 96-well plates and cultured to 90% confluency at 37°C for 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1), added to the culture medium, and cultivated 37°C for 1.5 h. The culture medium was subsequently replaced with new culture medium containing 1% FBS and BoScFv-PE38 (0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.50, 1.00, and 2.00 μM), and the MDBK cells were cultured for 24 h. The concentration of ATP was measured with a kit according to the manufacturer's instructions. **(B)** The titration of ammonia in BoHV-1-infected MDBK cells treated with BoScFv-PE38. MDBK cells were treated as described above, and the concentration of ammonia was measured with a kit. Statistical analyses were performed via analysis of variance (ANOVA) using SPSS software, version 18.0, and the fitting curves of various concentrations of BoScFv-PE38 were drawn with GraphPad Prism 5.0. "**" represents statistically significant differences ($P < 0.01$). "***" represents statistically significant differences ($P < 0.001$). **(C)** The TUNEL assay. MDBK cells were seeded in 24-well plates and cultured to 90% confluency at 37°C for 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1), added to the culture medium, and cultivated at 37°C for 1.5 h. The culture medium was subsequently replaced with new culture medium containing 1% FBS and BoScFv-PE38 (0.125, 0.25, 0.5, and 1 μM), and the MDBK cells were cultured for 24 h. The apoptosis-positive MDBK cells were detected with the *In Situ* Cell Death Detection Kit, AP (Roche) according to the manufacturer's instructions. The arrows indicate the TUNEL-positive apoptotic cells. **(D)** Apoptotic rates of BoHV-1-infected MDBK cells in the presence of different concentrations of BoScFv-PE38. The TUNEL-positive MDBK cells were counted, and statistical analyses were performed via analysis of variance (ANOVA) using SPSS software, version 18.0. The fitting curves of various concentrations of BoScFv-PE38 were drawn with GraphPad Prism 5.0. "***" represents statistically significant differences ($P < 0.001$). **(E)** Expression of apoptosis proteins in BoHV-1-infected MDBK cells after treatment with BoScFv-PE38. MDBK cells were seeded in six-well plates and cultured to 90% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1), added to the culture medium, and cultivated at 37°C for 1.5 h. The culture medium was subsequently replaced with new medium containing 1% FBS and BoScFv-PE38 (0.125, 0.25, 0.5, and 1 μM), and the MDBK cells were cultured for an additional 24 h. The whole-cell proteins were separated via SDS-PAGE and transferred to PVDF membranes. The membranes were subsequently hybridized with antibodies against PARP-1, Bcl-2, Bid, caspase-8, caspase-8 and β-actin, and the blots were developed using Super Signal chemoluminescent substrates. **(F)** Apoptosis proteins quantification of band intensity by Image J software, representative results are displayed with graphs corresponding to the ratios of cleaved PRAP-1, Bcl-2, Bid, caspase 8, caspase 9 and caspase 3 normalized to theβ-actin control conditions. The data were analyzed using SPSS software, and the graph was produced using GraphPad Prism 5.0. "n.s." represents no statistically significant differences ($P > 0.05$). "**" represents statistically significant differences ($P < 0.01$). "***" represents statistically significant differences ($P < 0.001$).

DISCUSSION

Immunotoxins are novel therapeutic agents that have previously been developed for use against human viruses and cancer. Currently, there is dearth of research into the development of immunotoxins for use against viruses of farmed animals, which have significant economic impacts worldwide (Liu et al., 2012). But it is necessary to develop new therapeutics protects those valuable cattle from infectious diseases; in particular the viral diseases to overcome limitations of vaccination approaches (Levings and Roth, 2013). Since recombinant immunotoxins represent a kind of therapeutics consisting of a cytotoxic agent fused to a variable antibody fragment; these agents bind specifically to target cells and exert cytotoxic effects (Berger and Pastan, 2010; Margolis et al., 2016), we therefore developed a novel recombinant BoScFv-PE38 immunotoxin that target BoHV-1 infected cells and block virus replication and decimation. The recombinant BoScFv-PE38 immunotoxin was generated by fusing the scFv fragment of an anti-BoHV-1 gD protein monoclonal antibody with the bacterial toxin PE38. To eliminate the non-specific cytotoxicity of normal cells, the natural binding domains of bacterial toxins are truncated and

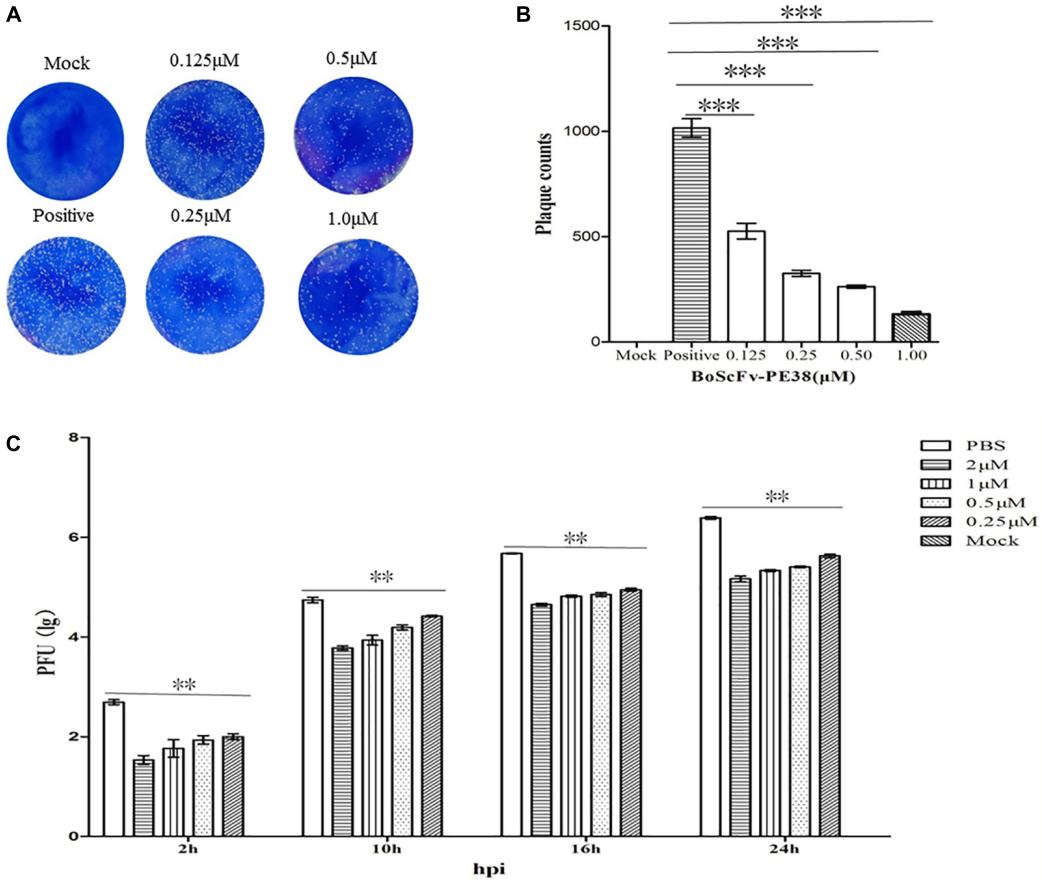


FIGURE 7 | The infectivity of BoHV-1 was inhibited by BoScFv-PE38. **(A,B)** Plaque reduction neutralization test of BoScFv-PE38. Mock represents normal MDBK cells; positive represents MDBK cells infected with BoHV-1 and treated with PBS. MDBK cells were infected with BoHV-1 and treated with different concentrations of BoScFv-PE38 (0.125, 0.25, 0.5, and 1.0 μM). Viral plaques were counted using 0.1% toluidine blue in saline. The data were analyzed using SPSS software, and the graph was produced using GraphPad Prism 5.0. "***" represents statistically significant differences ($P < 0.001$). **(C)** BoScFv-PE38 antiviral activity in ICAs. MDBK cells were seeded in six-well plates and cultured to 90% confluency at 37°C over 18–24 h. Then, the MDBK cells were infected with BoHV-1 (MOI = 1), added to the culture medium, and cultivated at 37°C for 1.5 h. The culture medium was subsequently replaced with new medium containing 1% FBS and BoScFv-PE38 (0.125, 0.25, 0.5, and 1 μM), and the MDBK cells were cultured for an additional 24 h. The whole culture was harvested and repetitively freeze-thawed three times, at 2, 10, 16, and 24 h. Then, virus titers were measured in plaque forming units (PFU), and viral plaques were counted visually. The data were analyzed using SPSS software. "**" represents statistically significant differences ($P < 0.01$).

connected with a flexible 16-aa linker peptide $(GGGGS)_3$, which has been reported as the best linker peptide (Chatterjee et al., 2012; Schotte et al., 2014). In this fusion protein, PE38 lacks its natural PE-binding domain to avoid an effect on non-infected cells, and scFv only target to the gD protein of BoHV-1 and exert its effect by killing and eradicating the pool of infected cells (**Figures 1A–C**). The results of SDS-PAGE analysis demonstrated that the recombinant protein was the desired immunotoxin, BoScFv-PE38 (**Figures 2A,B**).

Immunotoxins must exhibit high efficacy and specificity for binding to target viral antigens, together with inducing minimal side effects or toxicity to the non-infected cells, to be effective therapeutic agents (Brinkmann et al., 1993; Hanke et al., 2016; Spiess et al., 2016). To evaluate the ability of the BoScFv-PE38 immunotoxin to specifically bind to BoHV-1 envelope protein gD, the purified BoScFv-PE38 protein was labeled with horseradish peroxidase (HRP) and analyzed via western blotting and ELISA. Western blot analysis showed that the BoScFv-PE38 protein could specifically react with either the recombinant gD protein or BoHV-1 virus (**Figures 2C,D**). Furthermore, the BoScFv-PE38 immunotoxin bound BoHV-1 and gD with dissociation constants (Kd) of 97.63 nM and 12.81 nM, respectively (**Figures 2E,F**). The ability of BoScFv-PE38 immunotoxin to capture BoHV-1 was also demonstrated by a double sandwich ELISA (**Figure 2G**). These results indicated that the antibody portion of the recombinant immunotoxin possessed high affinity and specificity against the BoHV-1 virus. Since gD is the major antigen that induces neutralizing antibodies and is involved in virus penetration during BoHV-1 infection, gD has been considered the major target for antiviral agents and vaccines (Alves et al., 2014; Kumar et al., 2014). Therefore, it is expected that the gD protein is targeted by BoScFv-PE38 immunotoxin, not only because it is present in the viral envelope but also because it is expressed abundantly on the surface of

BoHV-1-infected cells (Liu et al., 2017; Müller et al., 2017). To confirm the binding specificity of BoScFv-PE38 immunotoxin, the gD protein was expressed in 293T cells and detected via IFA using BoScFv-PE38 as a primary antibody. Confocal imagining revealed that the location of the gD protein detected by BoScFv-PE38 was similar to that indicated by its GFP fusion protein, suggesting that the cells bound by BoScFv-PE38 immunotoxin were expressing gD (**Figures 3A,B**). Thus, the binding affinity of BoScFv-PE38 immunotoxin to gD is highly specific. We further evaluated the cytotoxicity of BoScFv-PE38 immunotoxin to 293T cells expressing gD. The results showed that BoScFv-PE38 immunotoxin could significantly reduce the proliferation of 293-T cells expressing gD through a dose-dependent manner, but there was no obvious deleterious effect to the untreated control cells (**Figures 3C,D**), suggesting that BoScFv-PE38 immunotoxin produced cytotoxic effects only to the cells expressing gD protein and negligible effect on uninfected cells. Since BoScFv-PE38 immunotoxin is composed of a targeting molecule scFv, the ability to recognize its corresponding antigen should be an inherent property (Ledford, 2011). In this study, we found that BoScFv-PE38 immunotoxin could specifically recognize BoHV-1 infected MDBK cells (**Figures 4A,B**). As an anti-BoHV-l therapeutics, it is necessary to eradicate viral infections. Thus, the immunotoxin must also have a capability to rapidly internalize the infected cells (Spiess et al., 2016). Internalization of BoScFv-PE38 immunotoxin was verified in this study, as it was found in the cytoplasm of MDBK cells infected with BoHV-1 (**Figure 4C**), which implied that the toxin PE38 can be efficiently delivered to the intracellular environment and play a crucial role in eliminating BoHV-1 (Frizzo da Silva et al., 2013). Confocal microscopy images provided further evidence that the nucleus of MDBK cells infected with BoHV-1 was damaged after treatment with 1.00 μM BoScFv-PE38 (**Figure 5D**). These findings clearly indicated the internalization capability of BoScFv-PE38 via entrapment of BoHV-1 after binding to gD, which is similar to the antitumor immunotoxin 4D5scFv-PE40 reported earlier (Sokolova et al., 2017).

The assessment of cytotoxic activity of BoScFv-PE38 immunotoxin via the MTS assay also revealed a minimal residual endotoxin effect on uninfected cells with study cytotoxic concentration up to 1.25 μM and the CC_{50} was 2.25 μM (**Figure 5A**). Next, the BoScFv-PE38 immunotoxin showed signification inhibition of BoHV-1 replication and exerted cytotoxic effects to the virus infected cells (**Figures 5B,C**). Previous studies have shown that cytotoxicity to targeted cells induces apoptosis in virus-infected host cells or cancer cells (Liu X.F. et al., 2016; Sokolova et al., 2017); specifically, high levels of ATP and ammonia are the major inducers of apoptosis (Jin et al., 2017; Zhang et al., 2017). In our study, BoScFv-PE38 significantly increased the titers of ATP and ammonia in MDBK cells infected with BoHV-1 (**Figures 6A,B**), which implied that apoptosis was triggered in MDBK cells as a result of treatment with BoScFv-PE38 immunotoxin and the effect was dose-dependent (**Figures 6C,D**). The analysis of pro and anti-apoptotic proteins including PRAP-1, Bcl-2, Bid, Caspase 8 and Caspase 3 also corroborated the specific apoptotic activity of BoScFv-PE38 leading to cytotoxic effect on BoHV-1 infected cells (**Figures 6E,F**).

Importantly, our results demonstrated that BoHV-1 virus production was significantly reduced in MDBK cells by treatment with BoScFv-PE38 immunotoxin (**Figures 7A–C**), whereas the IC_{50} value was 4.95 ± 0.33 nM; therefore, the SI was calculated as 456 ± 31 in terms of cytotoxic concentration (Supplementary Table S2). The results clearly demonstrated the inhibitory activity of BoScFv-PE38 immunotoxin for BoHV-1. Considering the neutralization activity of scFv in our previous study (Xu et al., 2017), the use of antibody portion scFv within BoScFv-PE38 immunotoxin could also target cell free virions during cytolytic phase of BoHV-1 infection in cattle. We therefore, conclude that our developed BoScFv-PE38 immunotoxin had an ability to target BoHV-1 at lytic phases of infection. Taken together, the findings of this study demonstrate that this recombinant immunotoxin could be a potential therapeutic agent for controlling and treating viral pathogens affecting at animals.

AUTHOR CONTRIBUTIONS

YL and JX: conceived and designed the experiments. JX, XL, BJ, XF, JW, YC, and XZ: performed the experiments. JX and YL: analyzed the data. XH: contributed reagents/ materials/analysis tools. JX and YL: wrote the paper. MI and JS: modified the paper.

FUNDING

This work was supported by a grant from the National Key Project of the Research and Development Program of China (Grant No. 2016YFD0500900), funding from the Beijing Innovation Team of Technology Systems in the Dairy Industry (Award No. bjcystx-ny-3), and the Special Program on Science and Technology Innovation Capacity Building of BAAFS (Award No. KJCX20170406).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.00653/full#supplementary-material

REFERENCES

Alves, D. L., Pereira, L. L. F., and van Drunen, L. H. S. (2014). Bovine herpesvirus glycoprotein D: a review of its structural characteristics and applications in vaccinology. *Vet. Res.* 45:111. doi: 10.1186/s13567-014-0111-x

Berger, E. A., and Pastan, I. (2010). Immunotoxin complementation of HAART to deplete persisting HIV-infected cell reservoirs. *PLoS Pathog.* 6:e1000803. doi: 10.1371/journal.ppat

Brinkmann, U., Lee, B. K., and Pastan, I. (1993). Recombinant immunotoxins containing the VH or VL domain of monoclonal antibody B3 fused to Pseudomonas exotoxin. *J. Immunol.* 150, 2774–2782.

Cai, Y., and Berger, E. A. (2011). An immunotoxin targeting the gH glycoprotein of KSHV for selective killing of cells in the lytic phase of infection. *Antivir. Res.* 90, 143–150. doi: 10.1016/j.antiviral.2011.03.175

Chandramohan, V., Pegram, C. N., Piao, H., Szafranski, S. E., Kuan, C. T., Pastan, I. H., et al. (2017). Production and quality control assessment of a GLP-grade immunotoxin, D2C7-(scdsFv)-PE38KDEL, for a phase I/II clinical trial. *Appl. Microbiol. Biotechnol.* 101, 2747–2766. doi: 10.1007/s00253-016-8063-x

Chatterjee, D., Chandran, B., and Berger, E. A. (2012). Selective killing of Kaposi's sarcoma-associated herpesvirus lytically infected cells with a recombinant immunotoxin targeting the viral gpK8.1A envelope glycoprotein. *MAbs* 4, 233–242. doi: 10.4161/mabs.4.2.19262

Cheng, C. H., Yang, F. F., Ling, R. Z., Liao, S. A., Miao, Y. T., Ye, C. X., et al. (2015). Effects of ammonia exposure on apoptosis, oxidative stress and immune response in pufferfish (*Takifugu obscurus*). *Aquat. Toxicol.* 164, 61–71. doi: 10.1016/j.aquatox.2015.04.004

Decker, T., Oelsner, M., Kreitman, R. J., Salvatore, G., Wang, Q. C., Pastan, I., et al. (2004). Induction of caspase-dependent programmed cell death in B-cell chronic lymphocytic leukemia by anti-CD22 immunotoxins. *Blood* 103, 2718–2726. doi: 10.1182/blood-2003-04-1317

Della Cristina, P., Castagna, M., Lombardi, A., Barison, E., Tagliabue, G., Ceriotti, A., et al. (2015). Systematic comparison of single-chain Fv antibody-fusion toxin constructs containing Pseudomonas Exotoxin A or saporin produced in different microbial expression systems. *Microb. Cell Fact.* 14:19. doi: 10.1186/s12934-015-0202-z

Denton, P. W., Long, J. M., Wietgrefe, S. W., Sykes, C., Spagnuolo, R. A., Snyder, O. D., et al. (2014). Targeted cytotoxic therapy kills persisting HIV infected cells during ART. *PLoS Pathog.* 10:e1003872. doi: 10.1371/journal.ppat.1003872

Frizzo da Silva, L., Kook, I., Doster, A., and Jones, C. (2013). Bovine herpesvirus 1 regulatory proteins bICP0 and VP16 are readily detected in trigeminal ganglionic neurons expressing the glucocorticoid receptor during the early stages of reactivation from latency. *J. Virol.* 87, 11214–11222. doi: 10.1128/JVI.01737-13

Geoghegan, E. M., Zhang, H., Desai, P. J., Biragyn, A., and Markham, R. B. (2015). Antiviral activity of a single-domain antibody immunotoxin binding to glycoprotein D of herpes simplex virus 2. *Antimicrob. Agents Chemother.* 59, 527–535. doi: 10.1128/AAC.03818-14

Hanke, L., Knockenhauer, K. E., Brewer, R. C., van Diest, E., Schmidt, F. I., Schwartz, T. U., et al. (2016). The antiviral mechanism of an influenza A virus nucleoprotein-specific single-domain antibody fragment. *mBio* 7:e01569-16. doi: 10.1128/mBio.01569-16

Hohensinner, P. J., Takacs, N., Kaun, C., Thaler, B., Krychtiuk, K. A., Pfaffenberger, S., et al. (2017). Urokinase plasminogen activator protects cardiac myocytes from oxidative damage and apoptosis via hOGG1 induction. *Apoptosis* 22, 1048–1055. doi: 10.1007/s10495-017-1388-9

Hou, S. C., Chen, H. S., Lin, H. W., Chao, W. T., Chen, Y. S., Fu, C. Y., et al. (2016). High throughput cytotoxicity screening of anti-HER2 immunotoxins conjugated with antibody fragments from phage-displayed synthetic antibody libraries. *Sci. Rep.* 6:31878. doi: 10.1038/srep31878

Jin, J., Wang, Y., Wu, Z., Hergazy, A., Lan, J., Zhao, L., et al. (2017). Transcriptomic analysis of liver from grass carp (*Ctenopharyngodon idellus*) exposed to high environmental ammonia reveals the activation of antioxidant and apoptosis pathways. *Fish Shellfish Immunol.* 63, 444–451. doi: 10.1016/j.fsi.2017.02.037

Keuser, V., Schynts, F., Detry, B., Collard, A., Robert, B., Vanderplasschen, A., et al. (2004). Improved antigenic methods for differential diagnosis of bovine, caprine, and cervine alphaherpes viruses related to *Bovine herpesvirus* 1. *J. Clin. Microbiol.* 42, 1228–1235. doi: 10.1128/JCM.42.3.1228-1235.2004

Khandelwal, N., Chander, Y., Rawat, K. D., Riyesh, T., Nishanth, C., Sharma, S., et al. (2017). Emetine inhibits replication of RNA and DNA viruses without generating drug-resistant virus variants. *Antivir. Res.* 144, 196–204. doi: 10.1016/j.antiviral.2017.06.006

Kumar, P., Ayalew, L. E., Godson, D. L., Gaba, A., Babiuk, L. A., and Tikoo, S. K. (2014). Mucosal immunization of calves with recombinant Bovine adenovirus-3 coexpressing truncated form of *Bovine herpesvirus*-1 gD and Bovine IL-6. *Vaccine* 32, 3300–3306. doi: 10.1016/j.vaccine.2014.03.073

Ledford, H. (2011). Toxic antibodies blitz tumours. *Nature* 476, 380–381. doi: 10.1038/476380a

Leist, M., Single, B., Castoldi, A. F., Kühnle, S., and Nicotera, P. (1997). Intracellular adenosine triphosphate (ATP) concentration: a switch in the decision between apoptosis and necrosis. *J. Exp. Med.* 185, 1481–1486. doi: 10.1084/jem.185.8.1481

Levings, R. L., Collins, J. K., Patterson, P. A., and Roth, J. A. (2015). Virus, strain, and epitope specificities of neutralizing Bovine monoclonal antibodies to *Bovine herpesvirus* 1 glycoproteins gB, gC, and gD, with sequence and molecular model analysis. *Vet. Immunol. Immunopathol.* 164, 179–193. doi: 10.1016/j.vetimm.2015.02.009

Levings, R. L., and Roth, J. A. (2013). Immunity to *Bovine herpesvirus* 1: I. Viral lifecycle and innate immunity. *Anim. Health Res. Rev.* 14, 88–102. doi: 10.1017/S1466252313000042

Lim, D., Kim, K. S., Kim, H., Ko, K. C., Song, J. J., Choi, J. H., et al. (2017). Anti-tumor activity of an immunotoxin (TGFα-PE38) delivered by attenuated *Salmonella typhimurium*. *Oncotarget* 8, 37550–37560. doi: 10.18632/oncotarget.17197

Liu, B., Zou, F., Lu, L., Chen, C., He, D., Zhang, X., et al. (2016). Chimeric antigen receptor T cells guided by the single-chain Fv of a broadly neutralizing antibody specifically and effectively eradicate virus reactivated from latency in CD4+ T lymphocytes isolated from HIV-1-infected individuals receiving suppressive combined antiretroviral therapy. *J. Virol.* 90, 9712–9724. doi: 10.1128/JVI.00852-16

Liu, W., Onda, M., Lee, B., Kreitman, R. J., Hassan, R., Xiang, L., et al. (2012). Recombinant immunotoxin engineered for low immunogenicity and antigenicity by identifying and silencing human B-cell epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11782–11787. doi: 10.1073/pnas.1209292109

Liu, X., Zhang, P., Rödl, W., Maier, K., Lächelt, U., and Wagner, E. (2017). Toward artificial immunotoxins: traceless reversible conjugation of RNase A with receptor targeting and endosomal escape domains. *Mol. Pharm.* 14, 1439–1449. doi: 10.1021/acs.molpharmaceut.6b00701

Liu, X. F., Xiang, L., Zhou, Q., Carralot, J. P., Prunotto, M., Niederfellner, G., et al. (2016). Actinomycin D enhances killing of cancer cells by immunotoxin RG7787 through activation of the extrinsic pathway of apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10666–10671. doi: 10.1073/pnas.1611481113

Mareeva, T., Wanjalla, C., Schnell, M. J., and Sykulev, Y. (2010). A novel composite immunotoxin that suppresses rabies virus production by the infected cells. *J. Immunol. Methods* 353, 78–86. doi: 10.1016/j.jim.2009.11.010

Margolis, D. M., Garcia, J. V., Hazuda, D. J., and Haynes, B. F. (2016). Latency reversal and viral clearance to cure HIV-1. *Science* 353:aaf6517. doi: 10.1126/science.aaf6517

Mazor, R., Vassall, A. N., Eberle, J. A., Beers, R., Weldon, J. E., Venzon, D. J., et al. (2012). Identification and elimination of an immunodominant T-cell epitope in recombinant immunotoxins based on *Pseudomonas* exotoxin A. *Proc. Natl. Acad. Sci. U.S.A.* 109, E3597–E3603. doi: 10.1073/pnas.1218138109

Müller, F., Stookey, S., Cunningham, T., and Pastan, I. (2017). Paclitaxel synergizes with exposure time adjusted CD22-targeting immunotoxins against B-cell malignancies. *Oncotarget* 8, 30644–30655. doi: 10.18632/oncotarget.16141

Muylkens, B., Thiry, J., Kirten, P., Schynts, F., and Thiry, E. (2007). *Bovine herpesvirus* 1 infection and infectious bovine rhinotracheitis. *Vet. Res.* 38, 181–209. doi: 10.1051/vetres:2006059

Polito, L., Mercatelli, D., Bortolotti, M., Maiello, S., Djemil, A., Battelli, M. G., et al. (2017). Two saporin-containing immunotoxins specific for CD20 and CD22 show different behavior in killing lymphoma cells. *Toxins* 9:E182. doi: 10.3390/toxins9060182

Raaperi, K., Orro, T., and Viltrop, A. (2014). Epidemiology and control of *Bovine herpesvirus* 1 infection in Europe. *Vet. J.* 201, 249–256. doi: 10.1016/j.tvjl.2014.05.040

Rola, J., Larska, M., Socha, W., Rola, J. G., Materniak, M., Urban-Chmiel, R., et al. (2017). Seroprevalence of *Bovine herpesvirus* 1 related alpha herpesvirus infections in free-living and captive cervids in Poland. *Vet. Microbiol.* 204, 77–83. doi: 10.1016/j.vetmic.2017.04.006

Schotte, L., Strauss, M., Thys, B., Halewyck, H., Filman, D. J., Bostina, M., et al. (2014). Mechanism of action and capsid-stabilizing properties of VHHs with an in vitro antipolioviral activity. *J. Virol.* 88, 4403–4413. doi: 10.1128/JVI.03402-3413

Sokolova, E., Guryev, E., Yudintsev, A., Vodeneev, V., Deyev, S., and Balalaeva, I. (2017). HER2-specific recombinant immunotoxin 4D5scFv-PE40 passes through retrograde trafficking route and forces cells to enter apoptosis. *Oncotarget* 8, 22048–22058. doi: 10.18632/oncotarget.15833

Spiess, K., Jakobsen, M. H., Kledal, T. N., and Rosenkilde, M. M. (2016). The future of antiviral immunotoxins. *J. Leukoc. Biol.* 99, 911–925. doi: 10.1189/jlb.2MR1015-468R

Spiess, K., Jeppesen, M. G., Malmgaard-Clausen, M., Krzywkowski, K., Kledal, T. N., and Rosenkilde, M. M. (2017). Novel chemokine-based immunotoxins for potent and selective targeting of cytomegalovirus infected cells. *J. Immunol. Res.* 2017:4069260. doi: 10.1155/2017/4069260

Thakur, V., Kumar, M., and Rathish, R. L. (2017). Seroprevalence of *Bovine herpesvirus*-1 antibodies in bovines in five districts of Uttarakhand. *Vet. World* 10, 140–143. doi: 10.14202/vetworld.2017.140-143

Wambaugh, M. A., Shakya, V. P. S., Lewis, A. J., Mulvey, M. A., and Brown, J. C. S. (2017). High-throughput identification and rational design of synergistic small-molecule pairs for combating and bypassing antibiotic resistance. *PLoS Biol.* 15:e2001644. doi: 10.1371/journal.pbio.2001644

Xu, J., Wu, J., Jiang, B., He, H., Zhang, X., Li, X., et al. (2017). Bovine single chain Fv antibody inhibits bovine herpesvirus-1 infectivity by targeting viral glycoprotein D. *Appl. Microbiol. Biotechnol.* 101, 8331–8344. doi: 10.1007/s00253-017-8566-0

Zhang, C., Liu, Z., Bunker, E., Ramirez, A., Lee, S., Peng, Y., et al. (2017). Sorafenib targets the mitochondrial electron transport chain complexes and ATP synthase to activate the PINK1-parkin pathway and modulate cellular drug response. *J. Biol. Chem.* 292, 15105–15120. doi: 10.1074/jbc.M117.783175

Zhu, L., Workman, A., and Jones, C. (2017). Potential role for a β-Catenin coactivator (high-mobility group AT-hook 1 protein) during the latency-reactivation cycle of *Bovine herpesvirus* 1. *J. Virol.* 91:e02132-16. doi: 10.1128/JVI.02132-16

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

The reviewer BFM and handling Editor declared their shared affiliation.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 September 2017
doi: 10.3389/fmicb.2017.01863

Depression of Vaccinal Immunity to Marek's Disease by Infection with Chicken Infectious Anemia Virus

*Yankun Zhang[1,2,3], Ning Cui[1,2,3,4], Ni Han[1,2,3], Jiayan Wu[1,2,3], Zhizhong Cui[1] and Shuai Su[1,2,3]**

[1] College of Veterinary Medicine, Shandong Agricultural University, Tai'an, China, [2] Shandong Provincial Key Laboratory of Animal Biotechnology and Disease Control and Prevention, Shandong Agricultural University, Tai'an, China, [3] Shandong Provincial Engineering Technology Research Center of Animal Disease Control and Prevention, Shandong Agricultural University, Tai'an, China, [4] Institute of Animal Husbandry and Veterinary, Shandong Academy of Agricultural Sciences, Jinan, China

Marek's disease (MD) has been occurring with increasing frequency in chickens in recent years. To our knowledge, however, there has been no report of the very virulent plus (vv+) MD virus (MDV) field isolate in China. Studies have shown that dual infection with immunosuppressive viruses such as chicken infectious anemia virus (CIAV) occurs frequently in chickens developing MD. In this study, we performed a designed set of *in vivo* experiments, which comprised five different groups of chickens, including the group of CVI988/Rispens-vaccinated chickens, the groups of CVI988/Rispens-vaccinated chickens infected with MDV or CIAV or both viruses (MDV and CIAV), and the group of MDV-challenged chickens. The effects of CIAV dual infection on the immunization of commercial MDV vaccine CVI988/Rispens were evaluated. The results show that infection of the SD15 strain of CIAV significantly reduced the weight and antibody titers to avian influenza virus (AIV)/Newcastle disease virus (NDV) inactivated vaccines of chickens immunized with the CVI988/Rispens, and resulted in the atrophy of thymus/bursa and the enlargement of spleen. The CVI988/Rispens vaccination conferred good immune protection for chickens challenged with 2000 PFU of the GX0101 strain of MDV. However, dual infection with SD15 significantly reduced the body weight, antibody titers induced by AIV/NDV inactivated vaccines and protective index of CVI988/Rispens, and resulted in the aggravation of the immunosuppression, mortality, and viremia of GX0101 in CVI988/Rispens-immunized/GX0101-challenged chickens. Overall, CIAV infection significantly reduced the protective effects of the CVI988/Rispens vaccine against MDV, implying that concurrent infection with CIAV may be a major contributor in the frequent attacks of MD in China in recent years.

Keywords: Marek's disease virus, infection, chicken infectious anemia virus, depression, vaccinal immunity

OPEN ACCESS

Edited by:
Jonatas Abrahao, Universidade Federal de Minas Gerais, Brazil

Reviewed by:
Jiabo Ding, China Institute of Veterinary Drug Control, China
Rodrigo Araújo Lima Rodrigues, Universidade Federal de Minas Gerais, Brazil

****Correspondence:***
Shuai Su
ssu6307@163.com

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

Received: *29 June 2017*
Accepted: *12 September 2017*
Published: *26 September 2017*

Citation:
Zhang Y, Cui N, Han N, Wu J, Cui Z and Su S (2017) Depression of Vaccinal Immunity to Marek's Disease by Infection with Chicken Infectious Anemia Virus. Front. Microbiol. 8:1863. doi: 10.3389/fmicb.2017.01863

INTRODUCTION

Marek's disease (MD) is a lymphoproliferative disease of chickens, which is caused by the MD virus (MDV) (Schat and Nair, 2008). MDVs are further divided into pathotypes, ranging from mild (m), virulent (v), and very virulent (vv) to very virulent plus (vv+) strains (Witter, 1997; Witter et al., 2005). MD is currently the only tumor disease in chickens that can be immunized

against by vaccine. After the first case of MD in 1960, HPRS-16/ATT (HPRS, Houghton Poultry Research Station), herpesvirus of turkeys (HVT), and HVT plus SB-1 or 301B/1 were developed to control MD (Churchill et al., 1969a,b; Okazaki et al., 1970; Witter et al., 1987). In the 1990s, CVI988/Rispens became the worldwide vaccine gold standard (Rispens et al., 1972). Recently, the "gold-standard" vaccine CVI988/Rispens has gradually showed poor protective efficacy against MDV in China (Teng et al., 2011; Tian et al., 2011; Cheng et al., 2012; Yu et al., 2013; Zhang et al., 2015; Cui et al., 2016). Several factors including the genetic background of chickens, the virulence of MDV, and concurrent infections with other immunosuppressive pathogens can influence the efficacy of MDV vaccines (Bacon et al., 2001). Although the use of vaccines may lead to an enhanced virulent strain of MDV, there has been no report of the vv+ MDV field isolate in China.

Concurrent infection with other viruses is very common in chickens with MD. This is particularly true of immunosuppressive viruses such as chicken infectious anemia virus (CIAV), avian reticuloendotheliosis virus (REV), and avian leukosis virus (ALV) (Qin et al., 2010; Zhao et al., 2012; Cui, 2013; Bao et al., 2015; Ahmed et al., 2016). CIA, which caused by CIAV, is characterized by aplastic anemia and immunosuppression in chickens (Miller and Schat, 2004). Chickens can be infected with CIAV, both vertically and horizontally (Hoop, 1992). CIAV is increasing in prevalence and infection increases susceptibility to a wide variety of other avian pathogens, presumably through immunosuppression of the CIAV-infected bird (Todd, 2004). Dual infection with CIAV and MDV showed synergistic effects on the pathogenicity with enhanced mortality and incidence of MD (Yang et al., 2010). Therefore, concurrent infection with CIAV is likely to be a factor in the increasingly frequent occurrences of MD in China in recent years. In this study, we analyzed the effects of CIAV dual infection on the immunization of commercial MDV vaccine CVI988/Rispens to better facilitate the establishment of effective control measures for MD in chickens.

MATERIALS AND METHODS

Ethics Statement

The study protocol and all animal studies were approved by the Shandong Agricultural University Animal Care and Use Committee (SACUC Permission number: AVM201701-2) and performed in accordance with the "Guidelines for Experimental Animals" of the Ministry of Science and Technology (Beijing, China). Any bird deemed to have reached the humane endpoint was culled.

Cell Culture and Viruses

Specific pathogen-free (SPF) chickens and chicken embryos used for the preparation of chicken embryo fibroblast (CEF) cultures were from SPAFAS Co. (Jinan, China). GX0101 strain of vv MDV and SD15 strain of CIAV were preserved in our laboratory (Zhang and Cui, 2005; Fang, 2017). MDV vaccine CVI988/Rispens was purchased from Merial Animal Health Co., Ltd.

Experimental Design

The experimental plan was illustrated in **Supplementary Figure S1**. Two-hundred SPF chickens were randomly divided into five equal groups (40 in each group) at 1 day old and reared separately in isolators with positive filtered air. All chickens of groups 1, 2, 3, and 4 were intra-abdominally (i.a.) infected at 1 day old with CVI988/Rispens. Groups 2 and 3 were inoculated intra-oral in addition with 400 EID_{50} of SD15 (Fang, 2017). Five days later, each chicken in groups 3, 4, and 5 was challenged i.a. with 2000 PFU of GX0101.

Measurement of Body Weight and Immune Organs Indices

The body weight of the chickens in different groups was measured at 0, 5, 9, 16, 23, 30, 37, and 44 days post-infection (dpi) with GX0101 to evaluate the effect of viral infection on growth rates. After 9 and 16 dpi, five chickens per group were used to evaluate the immune organs indices. The whole-body weight of each chicken was measured prior to euthanasia, and the thymus, spleen, and bursa from each chicken were collected and weighed. The immune organs indices were determined by the relative weight of the thymus, spleen, and bursa to the whole body.

Antibody Responses to Newcastle Disease Virus (NDV) and Avian Influenza Virus (AIV)–H9 Inactivated Vaccines

All chickens from each treatment group were vaccinated with Newcastle Disease Virus (NDV) and Avian Influenza Virus (AIV)–H9 inactivated vaccines according to the previously described procedure at 8 days old (Sun et al., 2007). On days 21, 28, and 35 post-vaccination, serum samples were randomly collected from chickens of each group. Hemagglutination inhibition (HI) antibody titers against NDV and AIV–H9 were determined in accordance with the routine procedures.

Protective Efficacy of CVI988/Rispens Vaccine

During 90 days post-challenges with GX0101, each dead chicken was recorded and necropsied. At the end of the study period, all surviving chickens were euthanized for autopsy. The protective efficacy of the vaccine for MD was expressed as a protective index (PI) calculated as the percentage of gross MD in non-vaccinated challenged control chickens minus the percentage of gross MD in vaccinated, challenged chickens divided by the percentage of gross MD in non-vaccinated challenged control chickens $\times$ 100.

Quantification of Viral Load

Blood samples in anticoagulants were collected from six chickens of each of the GX0101-infected groups (groups 3, 4, and 5) at 5, 9, 16, 23, and 30 dpi. DNA from peripheral blood lymphocytes (PBLs) were extracted using standard procedures (Sambrook et al., 1989). The MDV-specific primers were designed to be specific for the unique molecular marker of REV LTR in GX0101 (Duan et al., 2014). GX0101 DNA in PBLs was quantified with real-time quantitative PCR (RT-qPCR) according to the previous

method (Duan et al., 2014). qPCR reactions were set up on ice, and each reaction contained the following: MDV-specific primers (all at 0.5 uM), 10 ul SYBR Premix Ex Taq^{TM} (2×), 0.4 ul Rox Reference Dye II (50×), and 2 ul of DNA (approximately 100 ng). The reaction volume was brought up to 20 ul by the addition of ddH_2O. An ABI PRISM® 7500 Sequence Detection System (Applied Biosystems) was used to amplify and detect the reaction products.

Quantification of Cytokine mRNA Expression

Total RNA was extracted from PBLs collected from six chickens of each group (groups 3 and 4) at 0, 5, 9, 16, and 23 dpi with GX0101. The production of cytokine mRNA of interleukin-6 (IL-6), IL-18, and gamma interferon (IFN-γ) at different stages was quantified by RT-quantitative reverse transcription PCR (qPCR) according to the previous method (Jie et al., 2013; Heidari et al., 2016). Briefly, 2 μl of the oligo dT-based RT product from 4 μg of total RNA extracted from PBLs was used for each reaction. All the reactions were run in triplicates in an ABI PRISM® 7500 Sequence Detection System (Applied Biosystems). The amplification program was as follows: 95°C for 30 s, 40 cycles at 95°C for 5 s, 60°C for 34 s, followed by 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s. The relative expression ratios of target genes in the chickens of group 3 vs. those in group 4 were calculated by the $2^{-\Delta\Delta Ct}$ method using the chicken housekeeping gene β-actin as the endogenous reference gene in order to normalize the level of target gene expression.

Statistical Analysis

Statistical analysis was performed with the SPSS statistical software package for Windows, version 13.0 (SPSS Inc., Chicago, IL, United States). Differences between groups were examined for statistical significance by a two-tailed Student's *t*-test. A p-value <0.05 was considered statistically significant. Pairwise comparisons of the PI between vaccines were approximated using *Z*-statistic for difference between proportion data with Bonferroni corrections (Geng and Hills, 1989).

RESULTS

Body Weights

No significant differences were observed between different groups in the body weight of 6 days old chickens ($p > 0.05$) (**Figure 1**). At 5, 9, 16, 23, 30, 37, and 44 dpi with GX0101, there was no significant difference in body weight between group 1 and group 4 ($p > 0.05$), while that of the chickens in group 4 was significantly higher than those in group 5 ($p < 0.05$), indicating that CVI988/Rispens could prevent weight loss caused by GX0101 infection in SPF chickens. The body weight of chickens in group 1 was significantly increased as compared to that of the group 2 ($p < 0.05$), and the body weight of chickens in group 3 was significantly decreased as compared to that of group 4 ($p < 0.05$), suggesting that the body weight of chickens vaccinated with CVI988/Rispens, especially that of the CVI988/Rispens-vaccinated/GX0101-challenged chickens, was reduced by SD15 infection.

Immune Organs Indices

Chickens in group 5 exhibited an atrophied thymus and bursa of Fabricius with an enlarged spleen as compared to that of the chickens from group 1 ($p < 0.05$) after 9 and 16 dpi with GX0101 (**Table 1**). No significant change was observed in the chickens of group 4 ($p > 0.05$) with the exception of spleen enlargement presenting in chickens challenged with GX0101 at 9 dpi, indicating that CVI988/Rispens could reduce the damage of GX0101 to the immune organs in SPF chickens. Atrophy of thymus and bursa of Fabricius as well as spleen enlargement were noted in group 2 as compared to that of the chickens in group 1 ($p < 0.05$). Chickens in group 3 showed an atrophied thymus and bursa of Fabricius and enlarged spleen as compared to those of the chickens from group 4 ($p < 0.05$). These results demonstrated that SD15 infection significantly reduced the protective efficacy of CVI988/Rispens on immune organs in immunized chickens, especially those in the CVI988/Rispens-vaccinated/GX0101-challenged group (group 3).

Antibody Titers to AIV–H9 and NDV of Chickens in Different Groups

On 21, 28, and 35 days post-immunization with the inactivated vaccines, antibody titers to AIV–H9 and NDV in chickens from group 2 were significantly lower than that of the chickens from group 1, respectively ($p < 0.05$) (**Table 2**). Antibody titers to AIV–H9 of chickens from groups 3 and 5 were significantly decreased, and antibody titers to NDV were significantly decreased at 35 days post-immunization as compared to those of the chickens from group 4 ($p < 0.05$). The results indicated that SD15 led to immunosuppressive effects on humoral immune responses in the CVI988/Rispens-vaccinated chickens, especially on that of the CVI988/Rispens-vaccinated/GX0101-challenged chickens (group 3).

Protective Efficacy of CVI988/Rispens Vaccination Against Challenge of GX0101 in SPF Chickens

During the entire trial, chickens grew well, and no chickens died in the CVI988/Rispens-vaccinated group (**Figure 2**). The mortality rates of groups 2, 3, 4, and 5 were 14.3, 42.9, 5.7, and 31.4%, respectively (**Table 3**). In the GX0101-challenged groups, one chicken in group 3 and three chickens in group 5 developed visible tumor nodules, but no chicken developed visible MDV-induced lesion in group 4. CVI988/Rispens protected 94.3% of the chickens in group 4 while only protecting 54.3% of the chickens in group 3. These results indicate that the dual infection of SD15 significantly increased the GX0101-induced mortality rate and decreased the protective efficacy of the CVI988/Rispens vaccination.

Replication of GX0101 in SPF Chickens

Replication of MDV in the chickens of group 5 peaked at 23 days post-challenge with GX0101, while that in the

chickens of group 4 peaked at 16 dpi, with a significantly lower MDV copy number than that of the group 5 ($p < 0.05$) (**Figure 3**). This indicates that the CVI988/Rispens vaccine could significantly reduce the replication of GX0101. GX0101 increased continuously in chickens of group 3 and reached its peak at 30 dpi, with a significantly higher virus copy number than that of the group 4 ($p < 0.05$).

Cytokine mRNA Expression Levels

The expression of mRNA for IL-6 and INF-γ increased in chickens from group 3 while there was no significant difference in the expression of mRNA for IL-18 as compared with the values for group 4 in 6 days old chickens (**Figure 4**). The expressions of mRNA for IL-6, IL-18, and INF-γ increased significantly at 5 dpi in chickens from group 3, and then decreased to a level significantly lower than those of group 4 until 16 dpi.

DISCUSSION

Marek's disease infection has occurred with increasing frequency in chickens in recent years, but there has been no report concerning the isolation of the vv+ MDV field strain in China (Teng et al., 2011; Tian et al., 2011; Cheng et al., 2012; Yu et al., 2013; Zhang et al., 2015; Cui et al., 2016). China is rich

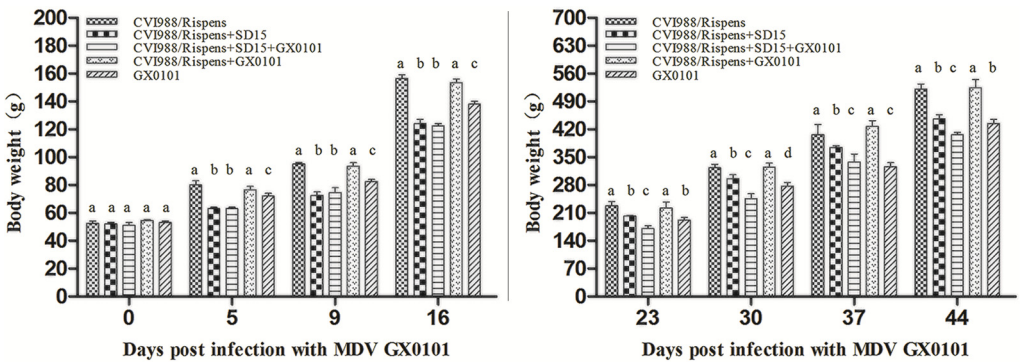


FIGURE 1 | The body weights of chickens in each group. The body weight of the chickens in different groups was measured at 0, 5, 9, 16, 23, 30, 37, and 44 days post-infection with GX0101 to evaluate the effect of virus infection on growth rates. [a,b,c,d]The different letters represent significant differences ($p < 0.05$). The same letters indicate the differences were not significant ($p > 0.05$).

TABLE 1 | The results of relative immune organs weight ($n = 5$).

Virus	Immune organs indices (9 dpi with GX0101)			Immune organs indices (16 dpi with GX0101)		
	Thy* (%)	Spl (%)	Bur (%)	Thy (%)	Spl (%)	Bur (%)
CVI988/Rispens	0.438 ± 0.037^{a}	0.150 ± 0.020^{a}	0.188 ± 0.034^{a}	0.533 ± 0.064^{a}	0.221 ± 0.017^{a}	0.395 ± 0.011^{a}
CVI988/Rispens+SD15	0.188 ± 0.012^{b}	0.241 ± 0.046^{b}	0.129 ± 0.049^{b}	0.240 ± 0.016^{b}	0.443 ± 0.032^{b}	0.299 ± 0.064^{b}
CVI988/Rispens+SD15+GX0101	0.177 ± 0.022^{b}	0.293 ± 0.023^{c}	0.107 ± 0.033^{b}	0.124 ± 0.022^{c}	0.689 ± 0.014^{c}	0.208 ± 0.054^{c}
CVI988/Rispens+GX0101	0.392 ± 0.040^{a}	0.198 ± 0.021^{d}	0.178 ± 0.021^{a}	0.503 ± 0.030^{a}	0.244 ± 0.041^{a}	0.386 ± 0.041^{a}
GX0101	0.254 ± 0.007^{c}	0.336 ± 0.023^{e}	0.138 ± 0.022^{b}	0.254 ± 0.007^{b}	0.488 ± 0.041^{b}	0.296 ± 0.052^{b}

The numbers in the table indicate the mean ± standard deviation. dpi, days post-infection. [a,b,c,d,e]*The different letters represent significant differences ($p < 0.05$). The same letters indicate the differences were not significant ($p > 0.05$). *Thy, relative thymus weight; Bur, relative bursa weight; Spl, relative spleen weight.*

TABLE 2 | The antibody response to vaccination with NDV and AIV–H9 inactivated vaccines.

Virus	HI antibody titers to AIV–H9			HI antibody titers to NDV		
	21 days	28 days	35 days	21 days	28 days	35 days
CVI988/Rispens	$6.3 \pm 0.89\ (30)^{a}$	$7.3 \pm 0.99\ (30)^{a}$	$7.4 \pm 0.93\ (30)^{a}$	$7.1 \pm 0.35\ (30)^{a}$	$7.5 \pm 0.77\ (30)^{a}$	$7.8 \pm 0.90\ (30)^{a}$
CVI988/Rispens+SD15	$5.3 \pm 0.46\ (25)^{b}$	$5.9 \pm 0.64\ (25)^{b}$	$6.3 \pm 0.46\ (25)^{b}$	$5.8 \pm 0.89\ (25)^{b}$	$6.1 \pm 0.52\ (25)^{b}$	$6.9 \pm 0.64\ (24)^{b}$
CVI988/Rispens+SD15+GX0101	$4.6 \pm 0.95\ (17)^{c}$	$4.3 \pm 1.17\ (16)^{c}$	$4.8 \pm 0.42\ (15)^{c}$	$5.6 \pm 0.79\ (17)^{b}$	$6.0 \pm 0.65\ (16)^{b}$	$5.8 \pm 0.71\ (15)^{c}$
CVI988/Rispens+GX0101	$5.9 \pm 0.99\ (30)^{ab}$	$6.9 \pm 0.64\ (29)^{a}$	$7.6 \pm 0.92\ (29)^{a}$	$6.4 \pm 0.52\ (30)^{b}$	$6.4 \pm 0.71\ (29)^{b}$	$7.5 \pm 0.52\ (29)^{a}$
GX0101	$4.3 \pm 0.71\ (24)^{c}$	$5.3 \pm 0.89\ (23)^{b}$	$5.6 \pm 0.87\ (21)^{b}$	$6.0 \pm 0.87\ (24)^{b}$	$6.4 \pm 0.92\ (23)^{b}$	$6.9 \pm 0.88\ (21)^{b}$

The numbers in the table indicate the mean ± standard deviation (sample size). [a,b,c]*The different letters represent significant differences ($p < 0.05$). The same letters indicate the differences were not significant ($p > 0.05$).*

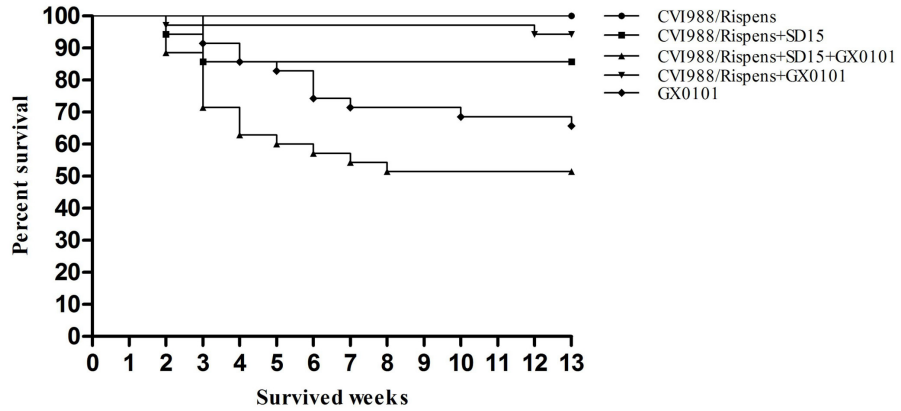


FIGURE 2 | Incidence of mortality in chickens inoculated with MDV GX0101. Chickens were inoculated with 2000 PFU of MDV GX0101 when they were 6 days old and were maintained in isolation for 13 weeks. During the experiment, all dead chickens were recorded and necropsied.

TABLE 3 | Protective efficacy of CVI988/Rispens against challenge of vv MDV GX0101 in SPF chickens.

Virus	Lesions	Mortality	Tumors rate	PI
CVI988/Rispens	–	–	–	–
CVI988/Rispens+SD15	14.3%	14.3%	0%	–
CVI988/Rispens+SD15+GX0101	45.7%	42.9%	2.9%	54.3%[a]
CVI988/Rispens+GX0101	5.7%	5.7%	0%	94.3%[b]
GX0101	100%	31.4%	8.6%	–

[a,b] *The different letters represent significant differences ($p < 0.05$).*

in genetic resources related to chickens, and various species of indigenous breeds scattered throughout the country. Long-term mixed breeding led to the dissemination of different viruses among chickens, especially CIAV, ALV, and REV (Qin et al., 2010; Zhao et al., 2012; Bao et al., 2015). The sub-clinical disease of commercial broilers due to CIAV is more common than clinical disease (McNulty, 1991). In chickens with an outbreak of MD, dual infection with MDV and other immunosuppressive viruses (and even triple infection) were detected (Cui, 2013). In the current study, we systematically evaluated the influence of CIAV infection on the immune efficacy of CVI988/Rispens in chickens.

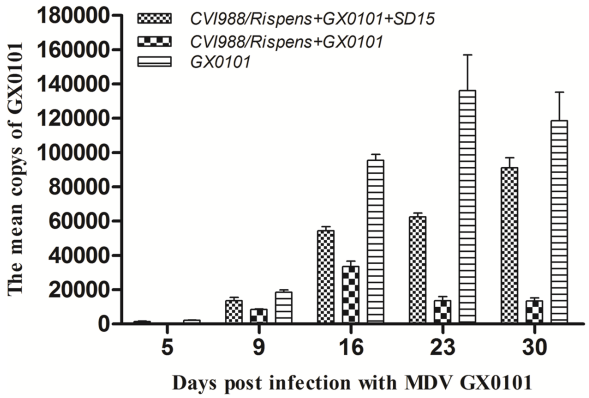


FIGURE 3 | Replication kinetics of MDV GX0101 in chickens. Replication kinetics of GX0101 viruses *in vivo* as determined by the viral genome copy numbers in the PBLs with real-time qPCR of the REV LTR fragment.

Our study shows that the dual infection of SD15 significantly reduced the body weight of chickens immunized with CVI988/Rispens and induced severe thymus/bursa atrophy and immunosuppression with significantly inhibited production of

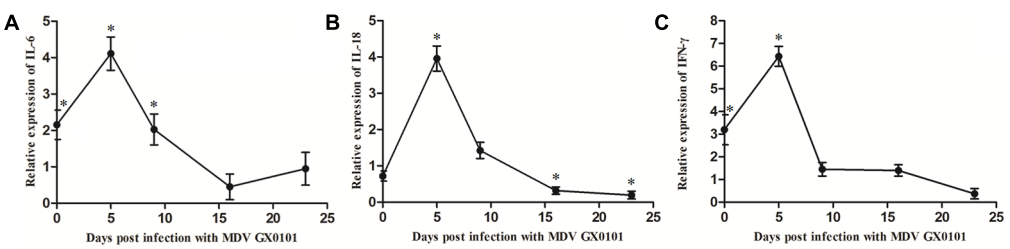


FIGURE 4 | Cytokine mRNA expression level in PBLs of the chickens. The relative expression levels of **(A)** IL-6, **(B)** IL-18, and **(C)** INF-γ genes in the chickens of group 3 vs. those in group 4 were determined by RT-qPCR. The chickens in group 3 were both vaccinated with CVI988/Rispens and infected with SD15 at 1 day old, and challenged with GX0101 at 5 days later. The chickens in group 4 were vaccinated with CVI988/Rispens at 1 day old and challenged with GX0101 at 5 days later. *Indicates significant difference ($p < 0.05$) between the two experimental groups.

antibodies to AIV/H9 and NDV inactivated vaccines (**Tables 1, 2** and **Figure 1**). A vaccinated model was then established using MDV-infected SPF chickens. The vv MDV GX0101 strain used for challenge is a recombinant field MDV that contains a REV LTR fragment (Cui et al., 2010; Su et al., 2012). The REV LTR was then selected as a molecular marker to differentiate CVI988/Rispens and to detect the multiplication level of GX0101. Our research demonstrates that CVI988/Ripens could provide good immunoprotection against challenge with 2000 PFU of GX0101 in SPF chickens at 6 days of age (**Table 3** and **Figure 2**). Replication of GX0101 as well as its pathogenicity in infected chickens was effectively decreased by CVI988/Rispens vaccination (**Table 3** and **Figures 2**, **3**). However, dual infection of SD15 significantly reduced the body weight and the antibody titers to AIV/NDV-inactivated vaccines in CVI988/Rispens-immunized/GX0101-challenged chickens while increasing the immunosuppression and mortality (**Tables 1, 2** and **Figure 1**). The PI of CVI988/Rispens against GX0101 challenge was also significantly decreased with increased viral titers of GX0101 in SPF chickens (**Table 3** and **Figure 3**). MDV vaccine has a protective effect in chickens but does not entirely prevent infection nor the replication of virulent virus. Our research and previous studies consistently demonstrated that the MDV vaccine with good immunogenicity could effectively inhibit the replication of wild strains of MDV. However, dual infection of CIAV poses a serious threat to the commercial CVI988/Rispens vaccine, causing considerable replication and long-term excreting of MDV in immunized chickens, which resulted in the enhanced transmission of MDV among chickens. Under the immune selective pressure, the virulence of field MDV showed a gradually increasing trend (Gimeno, 2008; Davison and Nair, 2014).

Cytokines play a critical role in driving immune response to MDV (Kaiser et al., 2003). Expressions of mRNA for IL-6 and INF-γ were increased significantly due to dual infection with CIAV in the chickens of the CVI988/Rispens group at 6 days old (**Figure 4**). Preliminary studies reported that IL-6 and IFN-γ mRNA transcript levels increased during early stages of infection with CIAV (Giotis et al., 2015). Expressions of mRNA for IL-6, IL-18, and INF-γ increased significantly and then decreased after 5 dpi with GX0101 in chickens of the CVI988/Rispens-vaccinated/SD15-inoculated group. IFN-γ plays a pivotal role in the early pathogenesis and immune responses to MDV infection (Xing and Schat, 2000; Abdul-Careem et al., 2007). It has been considered to be an immuno-modulator and vaccine adjuvant against MDV. Expression of recombinant chicken IFN-γ in HVT enhanced the protective efficacy of the vaccine against MDV and reduced the viral load and tumor incidence (Haq et al., 2011). IL-18 is a proinflammatory cytokine that induces IFN-γ production from CD4+T cells (Gobel et al., 2003). Thus, the reduced level of mRNA for IL-18 and IFN-γ in the late stage of infection probably correlates to the decline in the protective efficacy of the MDV vaccine. IL-6 is also a proinflammatory cytokine and its function in MDV infection is still unclear. The potential role for IL-6 in the immune response to MDV has been shown by a mouse model for another α-herpesvirus, herpes simplex virus-1. Mice showing an IL-6 deficiency when infected with HSV-1 have been shown to have increased viral titers and high mortality rates (Murphy et al., 2008). A similar IL-6 deficiency might also contribute to the increased titer of the MDV field strain and the depression of vaccinal immunity of the MD vaccine in chickens co-infected with CIAV.

CONCLUSION

Chickens concurrently infected with CIAV showed a declined immune efficacy of CVI988/Rispens against MD and a significantly enhanced susceptibility to MDV. Thus, CIAV might be a factor in frequent attacks of MD in chickens. In order to enhance the prevention and control of MD in chickens, detection of CIAV in chickens should be emphasized. However, no better measures are available for the control of CIAV (Cui, 2015). Most importantly, it is imperative that new vaccination strategies should be developed in case the currently available vaccines lose efficacy in controlling MDV strains with greater virulence (Lee et al., 2008; Su et al., 2015). Development of a recombinant MDV vector vaccine against CIAV is also a desirable application (Moeini et al., 2011; Reddy et al., 2016).

AUTHOR CONTRIBUTIONS

YZ collected and assembled, the data, did manuscript writing, and data analysis; NC and SS discussion, manuscript revision; NH and JW performed the animal experiments; SS and ZC concept and design, data analysis, manuscript revision, and final approval of the manuscript.

FUNDING

This study was supported by grants of the Key Program of NSFC-Henan Joint Found (U1604232), the National Natural Science Foundation of China (31402235), the National Key Research and Development Program of China (2017YFD0500700), the China Postdoctoral Science Foundation Funded Project (2016M592234), and the funds of Shandong "Double Tops" Program (SYL2017YSTD11).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01863/full#supplementary-material

FIGURE S1 | Flow chart of the experimental design.

REFERENCES

Abdul-Careem, M. F., Hunter, B. D., Parvizi, P., Haghighi, H. R., Thanthrige-Don, N., and Sharif, S. (2007). Cytokine gene expression patterns associated with immunization against Marek's disease in chickens. *Vaccine* 25, 424–432. doi: 10.1016/j.vaccine.2006.08.006

Ahmed, M. S., Ono, H., Sasaki, J., Ochiai, K., and Goryo, M. (2016). Persistence of chicken anemia virus antigen and inclusions in spontaneous cases of Marek's disease visceral lymphomas in broiler chickens at slaughterhouses. *J. Vet. Med. Sci.* 78, 825–829. doi: 10.1292/jvms.15-0615

Bacon, L. D., Hunt, H. D., and Cheng, H. H. (2001). Genetic resistance to Marek's disease. *Curr. Top. Microbiol. Immunol.* 255, 121–141.

Bao, K. Y., Zhang, Y. P., Zheng, H. W., Lv, H. C., Gao, Y. L., Wang, J. F., et al. (2015). Isolation and full-genome sequence of two reticuloendotheliosis virus strains from mixed infections with Marek's disease virus in China. *Virus Genes* 50, 418–424. doi: 10.1007/s11262-015-1191-z

Cheng, Y., Cong, F., Zhang, Y., Li, Z., Xu, N., Hou, G., et al. (2012). Genome sequence determination and analysis of a Chinese virulent strain, LMS, of Gallid herpesvirus type 2. *Virus Genes* 45, 56–62. doi: 10.1007/s11262-012-0739-4

Churchill, A. E., Chubb, R. C., and Baxendale, W. (1969a). The attenuation, with loss of oncogenicity, of the herpes-type virus of Marek's disease (strain HPRS-16) on passage in cell culture. *J. Gen. Virol.* 4, 557–564. doi: 10.1099/0022-1317-4-4-557

Churchill, A. E., Payne, L. N., and Chubb, R. C. (1969b). Immunization against Marek's disease using a live attenuated virus. *Nature* 221, 744–747. doi: 10.1038/221744a0

Cui, N., Su, S., Sun, P., Zhang, Y., Han, N., and Cui, Z. (2016). Isolation and pathogenic analysis of virulent Marek's disease virus field strain in China. *Poult. Sci.* 95, 1521–1528. doi: 10.3382/ps/pew073

Cui, Z. (2013). *Study on Viral Tumor of Chinese Chicken for its Prevention and Control*. Beijing: China Agricultural Press.

Cui, Z. (2015). *Avian Leukosis*. Beijing: China Agricultural Press.

Cui, Z., Zhuang, G., Xu, X., Sun, A., and Su, S. (2010). Molecular and biological characterization of a Marek's disease virus field strain with reticuloendotheliosis virus LTR insert. *Virus Genes* 40, 236–243. doi: 10.1007/s11262-009-0437-z

Davison, F., and Nair, V. (2014). Use of Marek's disease vaccines: could they be driving the virus to increasing virulence? *Expert Rev. Vaccines* 4, 77–88. doi: 10.1586/14760584.4.1.77

Duan, L., Su, S., Wang, Y., Li, S., Sun, P., Chen, W., et al. (2014). Comparison of immunoprotection between vaccination with meq-deleted Marek's disease virus and vaccine strain CVI988/Rispens. *Acta Microbiol. Sin.* 54, 1353–1361. doi: 10.13343/j.cnki.wsxb.2014.11.013

Fang, L. (2017). *Isolation, Pathogenicity and Bioinformatics Analysis of a Highly Pathogenic Chicken Infectious Anemia Virus*. Tai'an: Shandong Agricultural University.

Geng, S., and Hills, F. J. (1989). *Biometrics in Agriculture Science*. Dubuque, IA: Kendall/Hunt Publishing Company.

Gimeno, I. M. (2008). Marek's disease vaccines: a solution for today but a worry for tomorrow? *Vaccine* 26, C31–C41. doi: 10.1016/j.vaccine

Giotis, E. S., Rothwell, L., Scott, A. N., Hu, T., Talbot, R., Todd, D., et al. (2015). Transcriptomic profiling of virus-host cell interactions following chicken anaemia virus (CAV) infection in an in vivo model. *PLOS ONE* 10:e0134866. doi: 10.1371/journal.pone.0134866

Gobel, T. W., Schneider, K., Schaerer, B., Mejri, I., Puehler, F., Weigend, S., et al. (2003). IL-18 stimulates the proliferation and IFN-γ release of CD4+ T cells in the chicken: conservation of a Th1-like system in a nonmammalian species. *J. Immunol.* 171, 1809–1815. doi: 10.4049/jimmunol.171.4.1809

Haq, K., Elawadli, I., Parvizi, P., Mallick, A. I., Behboudi, S., and Sharif, S. (2011). Interferon-γ influences immunity elicited by vaccines against very virulent Marek's disease virus. *Antiviral Res.* 90, 218–226. doi: 10.1016/j.antiviral.2011.04.001

Heidari, M., Wang, D., Delekta, P. C., and Sun, S. (2016). Marek's disease virus immunosuppression alters host cellular responses and immune gene expression in the skin of infected chickens. *Vet. Immunol. Immunopathol.* 180, 21–28. doi: 10.1016/j.vetimm.2016.08.013

Hoop, R. K. (1992). Persistence and vertical transmission of chicken anaemia agent in experimentally infected laying hens. *Avian Pathol.* 21, 493–501. doi: 10.1080/03079459208418867

Jie, H., Lian, L., Qu, L. J., Zheng, J. X., Hou, Z. C., and Yang, N. (2013). Differential expression of Toll-like receptor genes in lymphoid tissues between Marek's disease virus-infected and noninfected chickens. *Poult. Sci.* 92, 645–654. doi: 10.3382/ps.2012-02747

Kaiser, P., Underwood, G., and Davison, F. (2003). Differential cytokine responses following Marek's disease virus infection of chickens differing in resistance to Marek's disease. *J. Virol.* 77, 762–768. doi: 10.1128/JVI.77.1.762-768.2003

Lee, L. F., Lupiani, B., Silva, R. F., Kung, H., and Reddy, S. M. (2008). Recombinant Marek's disease virus (MDV) lacking the Meq oncogene confers protection against challenge with a very virulent plus strain of MDV. *Vaccine* 26, 1887–1892. doi: 10.1016/j.vaccine.2008.01.046

McNulty, M. S. (1991). Chicken anaemia agent: a review. *Avian Pathol.* 20, 187–203. doi: 10.1080/03079459108418756

Miller, M. M., and Schat, K. A. (2004). Chicken infectious anemia virus: an example of the ultimate host-parasite relationship. *Avian Dis.* 48, 734–745. doi: 10.1637/7271-090304R

Moeini, H., Omar, A. R., Rahim, R. A., and Yusoff, K. (2011). Improving the potency of DNA vaccine against Chicken Anemia Virus (CAV) by fusing VP1 protein of CAV to Marek's Disease Virus (MDV) Type-1 VP22 protein. *Virol. J.* 8, 119–119. doi: 10.1089/jir.2007.0103

Murphy, E., Markdavis, J., Brown, A. S., Carmichael, M. D., Ghaffar, A., and Mayer, E. P. (2008). Effect of IL-6 Deficiency on susceptibility to HSV-1 respiratory infection and intrinsic macrophage antiviral resistance. *J. Interferon Cytokine Res.* 28, 589–596. doi: 10.1089/jir.2007.0103

Okazaki, W., Purchase, H. G., and Burmester, B. R. (1970). Protection against Marek's disease by vaccination with a herpesvirus of turkeys. *Avian Dis.* 14, 413–429. doi: 10.2307/1588488

Qin, L. T., Gao, Y. L., Pan, W., Deng, X. Y., Sun, F. F., Li, K., et al. (2010). Investigation of co-infection of ALV-J with REV, MDV, CAV in layer chicken flocks in some regions of China. *Chin. J. Prev. Vet. Med.* 32, 90–93.

Reddy, S. M., Izumiya, Y., and Lupiani, B. (2016). Marek's disease vaccines: current status, and strategies for improvement and development of vector vaccines. *Vet. Microbiol.* 206, 113–120. doi: 10.1016/j.vetmic.2016.11.024

Rispens, B. H., van Vloten, H., Mastenbroek, N., Maas, H. J., and Schat, K. A. (1972). Control of Marek's disease in the Netherlands. I. Isolation of an avirulent Marek's disease virus (strain CVI 988) and its use in laboratory vaccination trials. *Avian Dis.* 16, 108–125. doi: 10.1016/j.vetmic.2016.11.024

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor, NY: Cold spring harbor laboratory press.

Schat, K. A., and Nair, V. (2008). "Marek's disease," in *Diseases of Poultry*, 12th Edn, eds Y. M. Saif, A. M. Fadly, J. R. Glisson, L. R. McDougald, L. K. Nolan, and D. E. Swayne (Ames, IA: Black-well Publishing), 452–514.

Su, S., Cui, N., Cui, Z., Zhao, P., Li, Y., Ding, J., et al. (2012). Complete genome sequence of a recombinant Marek's disease virus field strain with one reticuloendotheliosis virus long terminal repeat insert. *Virol. J.* 86, 13818–13819. doi: 10.1128/JVI.02583-12

Su, S., Cui, N., Zhou, Y., Chen, Z., Li, Y., Ding, J., et al. (2015). A recombinant field strain of Marek's disease (MD) virus with reticuloendotheliosis virus long terminal repeat insert lacking the meq gene as a vaccine against MD. *Vaccine* 33, 596–603. doi: 10.1016/j.vaccine.2014.12.057

Sun, S., Cui, Z., Sun, A., and Zhu, X. (2007). Comparisons of preventive effects of maternal antibody on immunosuppression induced by homologous and heterologous reticuloendotheliosis viruses. *Chin. J. Anim. Vet. Sci.* 38, 488–492.

Teng, L. Q., Wei, P., Song, Z. B., He, J. J., and Cui, Z. Z. (2011). Molecular epidemiological investigation of Marek's disease virus from Guangxi, China. *Arch. Virol.* 156, 203–206. doi: 10.1007/s00705-010-0840-8

Tian, M., Zhao, Y., Lin, Y., Zou, N., Liu, C., Liu, P., et al. (2011). Comparative analysis of oncogenic genes revealed unique evolutionary features of field Marek's disease virus prevalent in recent years in China. *Virol. J.* 8:121. doi: 10.1186/1743-422X-8-121

Todd, D. (2004). Avian circovirus diseases: lessons for the study of PMWS. *Vet. Microbiol.* 98, 169–174. doi: 10.1016/j.vetmic.2003.10.010

Witter, R. L. (1997). Increased virulence of Marek's disease virus field isolates. *Avian Dis.* 41, 149–163. doi: 10.2307/1592455

Witter, R. L., Calnek, B. W., Buscaglia, C., Gimeno, I. M., and Schat, K. A. (2005). Classification of Marek's disease viruses according to pathotype: philosophy and methodology. *Avian Pathol.* 34, 75–90. doi: 10.1080/03079450500059255

Witter, R. L., Silva, R. F., and Lee, L. F. (1987). New serotype 2 and attenuated serotype 1 Marek's disease vaccine viruses: selected biological and molecular characteristics. *Avian Dis.* 31, 829–840. doi: 10.2307/1591039

Xing, Z., and Schat, K. A. (2000). Inhibitory effects of nitric oxide and gamma interferon on in vitro and in vivo replication of Marek's disease virus. *J. Virol.* 74, 3605–3612. doi: 10.1128/JVI.74.8.3605-3612.2000

Yang, M., Cui, Z. Z., Su, S., Zhang, H., and Wang, X. (2010). Synergic suppression effect of co-infection of chicken infectious anemia virus and Marek's disease virus in SPF CHICKEN. *Chin. J. Anim. Infect. Dis.* 4, 003.

Yu, Z. H., Teng, M., Luo, J., Wang, X. W., Ding, K., Yu, L. L., et al. (2013). Molecular characteristics and evolutionary analysis of field Marek's disease virus prevalent in vaccinated chicken flocks in recent years in China. *Virus Genes* 47, 282–291. doi: 10.1007/s11262-013-0942-y

Zhang, Y. P., Li, Z. J., Bao, K. Y., Lv, H. C., Gao, Y. L., Gao, H. L., et al. (2015). Pathogenic characteristics of Marek's disease virus field strains prevalent in China and the effectiveness of existing vaccines against them. *Vet. Microbiol.* 177, 62–68. doi: 10.1016/j.vetmic.2014.12.020

Zhang, Z., and Cui, Z. (2005). Isolation of recombinant field strains of Marek's disease virus integrated with reticuloendotheliosis virus genome fragments. *Sci. China C Life Sci.* 48, 81–88. doi: 10.1360/03yc0270

Zhao, P., Ma, C., Du, Y., and Cui, Z. (2012). Serological survey of the *Reticuloendotheliosis virus* infection in China native chicken flocks. *Pak. Vet. J.* 32, 621–623.

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

The reviewer RR and handling Editor declared their shared affiliation.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership