

## ARTICLE OPEN



# Incomplete tricarboxylic acid cycle and proton gradient in *Pandoravirus massiliensis*: is it still a virus?

Sarah Aherfi<sup>1,2,3,6</sup>, Djamal Brahim Belhaouari<sup>1,3,6</sup>, Lucile Pinault<sup>1,3</sup>, Jean-Pierre Baudoin<sup>1,3</sup>, Philippe Decloquement<sup>1,3</sup>, Jonatas Abrahao<sup>4</sup>, Philippe Colson<sup>1,2,3</sup>, Anthony Levasseur<sup>1,3</sup>, David C. Lamb<sup>5</sup>, Eric Chabriere<sup>1,3</sup>, Didier Raoult<sup>1,2,3</sup> and Bernard La Scola<sup>1,2,3</sup>✉

© The Author(s) 2021

The discovery of *Acanthamoeba polyphaga* Mimivirus, the first isolated giant virus of amoeba, challenged the historical hallmarks defining a virus. Giant virion sizes are known to reach up to 2.3 µm, making them visible by optical microscopy. Their large genome sizes of up to 2.5 Mb can encode proteins involved in the translation apparatus. We have investigated possible energy production in *Pandoravirus massiliensis*. Mitochondrial membrane markers allowed for the detection of a membrane potential in purified virions and this was enhanced by a regulator of the tricarboxylic acid cycle but abolished by the use of a depolarizing agent. Bioinformatics was employed to identify enzymes involved in virion proton gradient generation and this approach revealed that eight putative *P. massiliensis* proteins exhibited low sequence identities with known cellular enzymes involved in the universal tricarboxylic acid cycle. Further, all eight viral genes were transcribed during replication. The product of one of these genes, ORF132, was cloned and expressed in *Escherichia coli*, and shown to function as an isocitrate dehydrogenase, a key enzyme of the tricarboxylic acid cycle. Our findings show for the first time that a membrane potential can exist in Pandoraviruses, and this may be related to tricarboxylic acid cycle. The presence of a proton gradient in *P. massiliensis* makes this virus a form of life for which it is legitimate to ask the question “what is a virus?”.

The ISME Journal; <https://doi.org/10.1038/s41396-021-01117-3>

## INTRODUCTION

Since the discovery of *Acanthamoeba polyphaga* Mimivirus [1] in 2003, giant viruses of amoeba have consistently challenged the historical definition and classification of viruses [2]. With a virion size larger than 200 nm [1] and encompassing genome sizes greater than 250 kb [3], giant viruses differ from all previously described viruses to date. In 2013, *Pandoravirus salinus*, the first Pandoravirus isolated, broke all genomic viral size records with a genome size of 2.5 Mbp and 1-µm-diameter viral particles [4]. Moreover, Pandoravirus genomes to-date do not harbor any gene (s) encoding capsid protein(s), another hallmark of viral biology [4]. When observed by transmission electron microscopy (TEM), Pandoraviruses may utilize host cellulose production to build tegument [5]. Among later scientific discoveries causing giant viruses to challenge viral definition were the findings of associated virophages, which depicted for the first time a virus being infected by another virus [6]. In 2016, the MIMIVIRE system was identified as a functioning defense mechanism in Mimiviruses against invading virophages [7]. This was the first time that a mechanism for destroying invading foreign DNA, analogous to CRISPR-Cas system in bacteria, was observed to function in a virus. In 2018, the identification of Mimivirus proteins involved in protein translation again challenged another key feature of the definition

of viruses [8]. Subsequently, an almost complete set of protein translation apparatus was discovered in *Tupanvirus* and *Klosneuviruses* [9, 10]. Finally, in 2019 it was found that genes encoding multiple and unique cytochromes P450 monooxygenases commonly occur in giant viruses in the *Mimiviridae*, *Pandoraviridae*, and other families in the proposed order Megavirales [11, 12].

Recent data indicate that some giant viruses possess genes that may be involved in metabolic pathways such as fermentation, sphingolipid biosynthesis and nitrogen metabolism [13]. Moreover, tupanviruses harbor a gene coding for citrate synthase [14]. These viral genes are believed to be used to manipulate host metabolic pathways. However, no evidence has emerged to-date suggesting that such viruses use these gene products themselves for their own metabolic needs. As giant viruses have challenged nearly all criteria defining a virus, we decided to test another key hallmark of independent life in viruses: the ability to produce energy. To test this hypothesis, we used the giant virus *Pandoravirus massiliensis*, which we recently isolated [15]. The Pandoravirus family stands uniquely apart from other giant viruses of amoebas because of their huge gene content, with more than 80% of their predicted gene products being ORFans (no homologs in international protein databases). Hence, *P. massiliensis* provides a novel viral system for the discovery of genes with currently

<sup>1</sup>Aix Marseille Univ, IRD, MEPIH, Marseille, France. <sup>2</sup>Assistance Publique—Hôpitaux de Marseille (AP-HM), Marseille, France. <sup>3</sup>Institut Hospitalo-Universitaire (IHU)—Méditerranée Infection, Marseille, France. <sup>4</sup>Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo, Horizonte, Brazil. <sup>5</sup>Faculty of Health and Life Sciences, Swansea University, Swansea, UK. <sup>6</sup>These authors contributed equally: Sarah Aherfi, Djamal Brahim Belhaouari. ✉email: [bernard.la-scola@univ-amu.fr](mailto:bernard.la-scola@univ-amu.fr)

Received: 2 February 2021 Revised: 24 August 2021 Accepted: 10 September 2021

Published online: 23 September 2021

unknown functions. In the living world, energy generation is mostly associated with the creation of proton gradients. Thus, we searched for energy gradients in *P. massiliensis*. We were able to observe the presence of a proton gradient in this virus, and surprisingly, it was mainly present in the mature particles. We then searched for genes that could be responsible for this proton gradient. No genes involved in the respiratory chain or with identity to ATP synthase were detected. However, gene products having weak identities with nearly all enzymes of the tricarboxylic acid (TCA) cycle were observed. These viral genes were transcribed together, and the product of at least one gene, isocitrate dehydrogenase (IDH), was functional. These findings argue that *P. massiliensis* is a viral life form for which it is now legitimate to ask the question: "What is a virus?"

## MATERIALS AND METHODS

### *P. massiliensis* immunofluorescence staining using mouse specific polyclonal antibodies

In the aim to avoid any confusion with the staining of the amoeba mitochondria, mouse polyclonal antibodies were generated to specifically stain the isolated *P. massiliensis* mature virions and their replication cycle into amoeba cells. We first immunized a mouse with *P. massiliensis* by the subcutaneous route. After three inoculations, mouse serum containing polyclonal antibodies specific to *P. massiliensis* was collected and adsorbed on uninfected *A. castellanii* lysate to remove nonspecific antibodies targeting amoeba [16]. To permeabilize the cell membranes and saturate the nonspecific binding sites, cells were incubated in fetal calf serum with 1% (v/v) Triton X-100 in phosphate-buffered saline (PBS) for 1 h. Infected amoebas were incubated overnight at 28 °C in a humidified chamber with anti-Pandoravirus antibodies. Subsequently, the samples were washed three times with 0.1% (v/v) Triton X-100 in PBS. Each sample was incubated with fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (Immunotech, Marseille, France) for 60 min at 28 °C in a humidified chamber and finally washed three times with PBS.

### Assessment of a membrane electrochemical gradient in *P. massiliensis* virions

The membrane electrochemical gradient in *P. massiliensis* was assessed by two dyes: Mitotracker Deep Red 633 (Invitrogen, Carlsbad, California, USA), and tetramethyl rhodamine (TMRM) reagent (Thermo Fisher Scientific). Mitotracker Deep Red 633 and TMRM are cationic fluorescent dyes that are selectively sequestered by mitochondria in live cells based on their charge and that are dependent on the mitochondrial membrane potential for loading.

(i) To assess the membrane electrochemical gradient of *P. massiliensis* during the replication cycle, amoeba were infected with *P. massiliensis* at a multiplicity of infection of ten in IBIDI petri  $\mu$ -dishes previously coated with poly-L lysine to retain adherent cells. Each sample was dyed with MitoTracker Deep Red 633 for 45 min at 37 °C. The samples were washed twice with PBS to remove the excess fluorescent dye and blocked with acetone at different time points of the viral cycle (2, 4, 6, 8, 10, 12, 14, and 16 h), and then washed three times with PBS. To avoid any further confusion with staining of the amoeba mitochondria with MitoTracker Deep Red 633, immunofluorescence staining using mouse specific polyclonal antibodies targeting *P. massiliensis* inside the amoeba was also undertaken following the protocol described above.

(ii) To assess the membrane electrochemical gradient in the purified *P. massiliensis* virions, both MitoTracker Deep Red 633 and TMRM were used.

Briefly, lysed amoeba infected by *P. massiliensis* were filtered (5  $\mu$ m) then centrifuged 10 min at 500  $\times g$ , and cellular debris was discarded. The viral supernatant was centrifuged 20 min at 6800  $\times g$ , and the pellet was resuspended twice in PAS. The second time, the viral pellet was resuspended in survival buffer. For MitoTracker Deep Red 633 staining, purified viral particles were incubated with the dye for 45 min at 37 °C. Subsequently, the sample was washed twice with PBS to remove the excess fluorescent dye and fixed with acetone. Immunofluorescence staining using mouse specific polyclonal antibodies targeting *Pandoravirus massiliensis* was also carried out.

For TMRM, 1 mL of purified viral particles was incubated with 1  $\mu$ L of stock solution of TMRM (100  $\mu$ M) for 30 min at 30 °C in IBIDI petri  $\mu$ -dishes. In control experiments, sample cultures of *Staphylococcus aureus* were

used as positive controls, and viral supernatant from *cowpoxvirus* cultured on Vero (ATCC CCL-81) African green monkey kidney cells [17] was used as negative control.

To ensure the specificity of Mitotracker fixation on virions alone and to avoid misleading observations, in addition to immunofluorescence staining using mouse specific polyclonal antibodies targeting *P. massiliensis*, we also checked for the absence of amoebal structure debris including mitochondria by scanning electron microscopy with the SU5000 device (Hitachi High-Technologies, Tokyo, Japan). Scanning electron microscopy was parallelly carried out on both purified virions and on amoebal isolated mitochondria to exclude any possibility of morphological confusion between these two structures [18].

### Assessment of the effect of the decoupling agent CCCP on viral particles

To determine that fluorescence observed with the two mitochondrial dyes was specifically linked to the existence of an electrochemical gradient throughout the virions tegument, we used carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) (Sigma-Aldrich C2759; Saint-Louis, Missouri, USA), an inhibitor of oxidative phosphorylation that acts by dissipating the electrochemical gradient induced by the proton concentration. CCCP effect was assessed on *P. massiliensis* virions treated with TMRM. CCCP reagent (100, 200, 300, 400  $\mu$ M) was directly added in tubes containing  $\approx 10^7$  viral particles/mL. Negative control were *P. massiliensis* virions without CCCP. Samples were incubated at 35 °C overnight and then transferred to petri  $\mu$ -dishes. Next, 1  $\mu$ L of a stock solution of TMRM (100  $\mu$ M) was added and plates incubated 30 min at 30 °C. Images were acquired by confocal microscopy using a Zeiss LSM 800 microscope. To assess if the CCCP could have an impact on the infectiveness of virions, the infectivity of *P. massiliensis* particles was assessed before and after incubation with CCCP by calculating the TCID<sub>50</sub> using the method of Reed and Muench [19]. The potential impact of CCCP on viral replication cycle was also assessed by immunofluorescence and qPCR. *A. castellanii* strain Neff cells were inoculated with viral particles previously incubated with 400  $\mu$ M CCCP and washed three times. As at the end of the replication cycle, no differences were observed, so we focused on early time points of the viral cycle. Infected amoeba were collected 45 min (H0) and 3 h (H3) postinfection. *P. massiliensis* particles spotted on slides were labeled with anti-*P. massiliensis* specific antibodies according to the protocol described above. qPCR was carried out using isolated DNA from the collected cells targeting the DNA polymerase gene of *P. massiliensis* by PCR (forward primer: 5'-ATGGCGCCCGTCTGGAAG; reverse primer: 5'-GGCGCCAAAGTGGTGCGA). qPCR was performed with a LightCycler 480 SYBR Green 1 Master reaction mix (Roche Diagnostics, Mannheim, Germany), following the manufacturer's temperature program with 60 °C for the primer hybridization and elongation temperature.

### Assessment of the effect of acetyl-CoA on *P. massiliensis* viral particles

To assess if the potential electrochemical gradient observed in virions could be modified, we chose to study the impact of the acetyl-CoA, a well-known regulator of the TCA cycle. Acetyl-CoA (Sigma-Aldrich) was added at concentrations of 0.8 mM, 0.4 mM, 0.2 mM, 0.1 mM, 0.01 mM, and 0.001 mM to 1 mL of the viral suspension of purified particles ( $\approx 10^8$  particles/mL). The negative control consisted in viral particles without acetyl-CoA. The samples were incubated at 30 °C for 24 h. After incubation, the samples were transferred into petri  $\mu$ -dishes and stained with TMRM by following the above-described protocol. Images were acquired using a LSM 800 confocal microscope. Image processing and fluorescence intensity evaluations were conducted using Zen Bleu software.

### Bioinformatics and phylogenetics analyses

The *P. massiliensis* genome was analyzed by BLASTp analyses against the GenBank nr database version 2.10.0+ using an e-value threshold of  $1 \times 10^{-2}$ . The search for gene products potentially involved in energy metabolism was performed by delta-BLAST analyses against the Conserved Domain Database (CDD) [20, 21]. For some predicted ORFs having hits with low similarity, PSI Blast, HMM searches against PFAM database (<http://hmmer.org>; HmmerWeb version 2.41.1), HHPRED (version of the database: 20200717) analyses [22] and structure prediction using the PHYRE2v2.0 server were performed [23]. Orthologs in other pandoravirus genomes were searched using the ProteinOrtho tool with a 30% identity percentage threshold and 50% as a coverage percentage threshold [24].

Gene products of all Pandoraviruses were also analyzed by BLASTp against the COG database version 2003 COGs, 2014 update [25, 26]. The viral ORFs harboring a hit against class C of the COG (energy metabolism) with a bitscore >50 were considered significant.

Sequences were aligned with Muscle software [27]. Phylogenetic tree was built by using FastTree v2.1.10 [28] with the maximum likelihood method and formatted with the iTOL v6.1.1 program [29]. The robustness of the groupings was assessed by bootstrap resampling of 1000 replicates and only significant bootstrap values (>70%) were displayed.

### Transcriptome sequencing (RNA-seq) on *P. massiliensis*

The transcriptome of *P. massiliensis* strain BZ81 was assessed as described previously using amoebas infected by *P. massiliensis* [15] as well as freshly released mature viral particles. Mature virions were collected 11 h following amoeba inoculation, passed through 5- $\mu$ m-pore filters, and centrifuged at 500  $\times$  g for 10 min to remove all amoeba debris.

### qRT-PCR of suspected *P. massiliensis* TCA cycle genes

Viral DNA was isolated from 200  $\mu$ L of viral supernatant from *P. massiliensis* culture using the EZ1 tissue kit (Qiagen, France) according to the manufacturer's recommendations. Before RNA extraction, all samples were incubated with Turbo DNase for 30 min at 37 °C. RNA was extracted using the RNeasy mini kit (Qiagen, France) at different time points of the *P. massiliensis* replication cycle, from H0 (i.e., 45 min after infection of amoeba cells by viral particles) until H16 postinfection (release of neo-synthesized virions), according to the manufacturer's recommendations). Subsequently, three further digestions with Turbo DNase (Invitrogen, United States) were performed for each RNA sample to eliminate DNA contamination. Finally, enzyme inactivation reagent was added and incubated at room temperature for 5 min. The samples were then centrifuged at 10,000  $\times$  g for 1.5 min. Two PCR systems targeting the *P. massiliensis* DNA polymerase gene (forward primer: 5'-ATGGCGCCCGTCTGGAAG; reverse primer: 5'-GGCGCCAAAGTGGTGCGA) and the housekeeping gene of the RNA polymerase of *A. castellanii* (forward: 5'-ACGAACCTCCGAGAGATGCA; reverse: 5'-CACCTTGACCAGTCCCTTCT) were used to check for genomic DNA contamination.

Nucleotide primers targeting the selected ORFs of the *P. massiliensis* gene sequences were designed using the primer3 tool [30] (Supplementary File 1). RT-qPCR was performed in a one-step reaction using the QuantiTect SYBR Green RT-PCR Kit (Qiagen, France) following the manufacturer's recommendations. Each experiment was performed in triplicate. The results were considered positive if the cycle threshold obtained in three replicates was less than 35.

### Proteome analysis of *P. massiliensis*

Protein extraction was carried out on purified viral particles and amoebas infected with *P. massiliensis* at different time points of the replication cycle (H0 to H16). Samples were treated as previously described [15, 31]. The final analysis was performed with an internal protein sequence database built primarily with two types of amino acid sequences as follows:

(i) Sequences obtained by translating the whole genome into the six reading frames and then fragmenting the six translation products into 250-amino-acid-long sequences with a sliding step of 30 amino acids, stop codons being replaced by a trypsin cleavage site. Contiguous sequences that were positive for peptide detection were fused and reanalyzed. (ii) Host organism sequences (*A. castellanii* str.Neff, Uniprot, 15235 sequences). The virus sequences thus obtained were finally compared with the ORFs.

### Cloning, expression and purification of predicted *P. massiliensis* TCA cycle enzymes

*P. massiliensis* genes encoding the predicted TCA enzymes (ORFs 132, 181, 206, 577, 595, 762, 864, and 1245) were designed to include a Strep-tag at the N-terminus and optimized for *Escherichia coli* expression. Genes were synthesized by GenScript (Piscataway, NJ, USA) and ligated between the *Nde*I and *Not*I sites of a pET22b(+) plasmid. Competent BL21(DE3) cells grown in autoinducing ZYP-5052 medium were used for expression of the recombinant proteins. To produce each protein, the culture was shaken at 37 °C until an O.D. 600 nm of 0.6 was reached, after which the temperature was lowered to 20 °C for 20 h. Cells were harvested by centrifugation (5000  $\times$  g, 30 min, 4 °C), and the resulting pellet was resuspended in 50 mM Tris pH 8, 300 mM NaCl and stored at -80 °C overnight. The crude extract was thawed and incubated on ice for 1 h following the addition of lysozyme, DNase I and phenylmethylsulfonyl fluoride (PMSF) to final

concentrations of 0.25 mg/mL, 10  $\mu$ g/mL and 0.1 mM, respectively. Partially lysed cells were sonicated using a Q700 sonicator system (QSonica), and cell debris was removed following a centrifugation step (12,000  $\times$  g, 20 min, 4 °C). Proteins were purified with an ÄKTA avant system (GE Healthcare) using strep-tag affinity chromatography (Wash buffer: 50 mM Tris pH 8, 300 mM NaCl and Elution buffer: 50 mM Tris pH 8, 300 mM NaCl, 2.5 mM desthiobiotin) on a 5-mL StrepTrap HP column (GE Healthcare). Recombinant protein expression was confirmed by MALDI-TOF MS analysis of excised gel bands previously isolated by SDS-PAGE. Protein concentrations were measured using a Nanodrop 2000c spectrophotometer (Thermo Scientific).

### Enzymatic activity assay and kinetics

Enzymatic activities of the predicted citrate synthase, aconitase and  $\alpha$ -ketoglutarate dehydrogenase were tested using the Citrate Synthase Assay kit, Aconitase Activity Assay kit and  $\alpha$ -ketoglutarate Dehydrogenase Activity Colorimetric Assay kit, respectively (Sigma-Aldrich, St. Louis, MS, USA). IDH activity assays were performed using the isocitrate dehydrogenase activity assay kit (Sigma-Aldrich, St. Louis, MS, USA) and monitored with a Synergy HT microplate reader (BioTek, Winooski, VT, USA). Reactions were carried out in duplicate at 37 °C in a 96-well plate containing a final volume of 100  $\mu$ L in each well. Conversion of the isocitrate to  $\alpha$ -ketoglutarate was monitored for 30 min following changes in absorbance at 450 nm, corresponding to the production of NADH. A NADH standard curve was constructed, allowing enzyme quantification and calculation of specific activity. In addition, IDH activity was assessed on mature viral particles purified as described above, and on human IDH (Sigma-Aldrich; St. Louis, MS, USA). Initial velocities were calculated using Gen5.1 software (BioTek), and the obtained mean values were fitted using the Michaelis-Menten equation in Prism 6 (GraphPad Software, San Diego, CA, USA).

### Statistical analysis

Statistical analysis was performed using GraphPad Prism for Windows. Statistical differences were evaluated by one-way ANOVA. Statistical significance was set at  $p < 0.05$ .

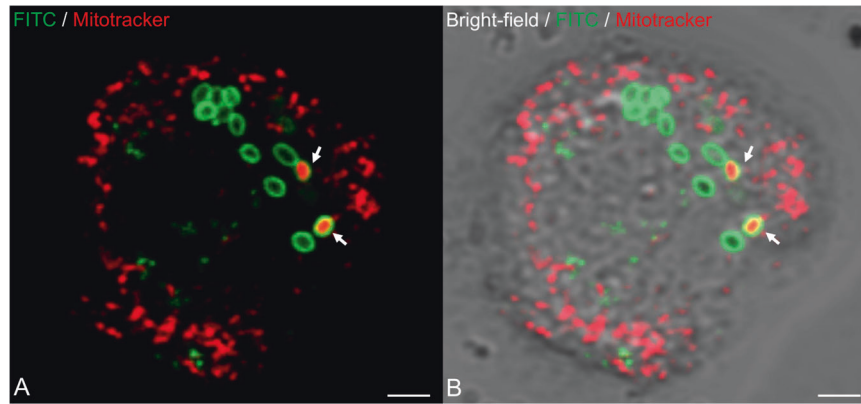
## RESULTS AND DISCUSSION

### Two mitochondrial dyes, MitoTracker and TMRM, identify a membrane potential in mature *P. massiliensis* virions and during its replication cycle

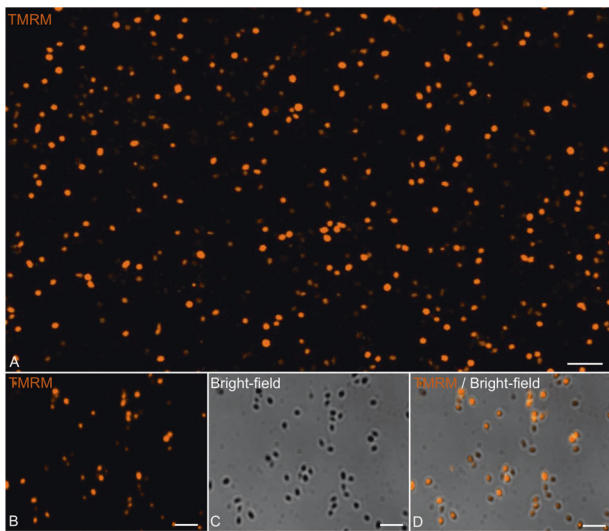
Membrane potential in *P. massiliensis* virions was assessed during the replication cycle in *A. castellanii* and mature virions freshly released from amoebas. During the *P. massiliensis* viral cycle in *A. castellanii*, a variable proportion of MitoTracker-labeled viruses was observed (Fig. 1). The specificity of the labeling was ensured by the co-localization of FITC-conjugated anti-*P. massiliensis* antibodies.

An analogous experiment performed using mature viral particles showed that approximately 20% of the total number of particles were labeled both by anti-*P. massiliensis* specific antibodies and MitoTracker Deep Red 633 (Supplementary Fig. 1). Moreover, the purified viral particles were also labeled by TMRM staining (under the TRITC wavelength (532 nm), with a fluorescent signal (Fig. 2, Supplementary Fig. 2), similar to the results obtained for the *S. aureus* positive control. No fluorescence was observed in the cowpoxvirus negative control experiments. We identified here a virion membrane potential for the first time using two different fluorescent mitochondrial dyes, MitoTracker Deep Red 633 and TMRM, which allowed the detection of a fluorescent signal in mature *P. massiliensis* virions. To note, TMRM has been scientifically acknowledged as the ideal marker to assess mitochondrial membrane potential [32]. It can be argued that the fluorescent signal of the TMRM might be confused with the staining of remaining membrane debris of amoebal mitochondria, since all the enzymes of *A. castellanii*, involved in energy production, are localized in mitochondria [33–35]. Therefore, in addition to the use of specific antibodies targeting the virus and after careful separation of Pandoravirus particles from amoebal debris, we





**Fig. 1** Confocal imaging of amoeba infected by *P. massiliensis* stained with MitoTracker Deep Red (in red) and with specific anti-*P. massiliensis* antibodies (in green). **A** Co-localization of the MitoTracker signal (in red) with virus marked by specific antibodies (FITC) (arrows). **B** Merge of Bright-field, FITC and MitoTracker fluorescence. The scales bar correspond to 2  $\mu$ m.



**Fig. 2** Confocal imaging of TMRM staining of purified *P. massiliensis* virions. **A**, **B** Viral mature particles stained with TMRM. **C** Bright-field channel. **D** Merge of Bright-field and TMRM showing the internalization of the TMRM signal in viral particles. The scales bar correspond to 5  $\mu$ m.

checked for the absence of any amoebal structure by 3D analysis (Supplementary Fig. 3). Furthermore, a negative control consisting of purified amoebal mitochondria clearly showed a distinct morphology from those of pandoraviruses virions, thus excluding any misleading conclusions of potential staining by MitoTracker of amoebal debris (Supplementary Fig. 4).

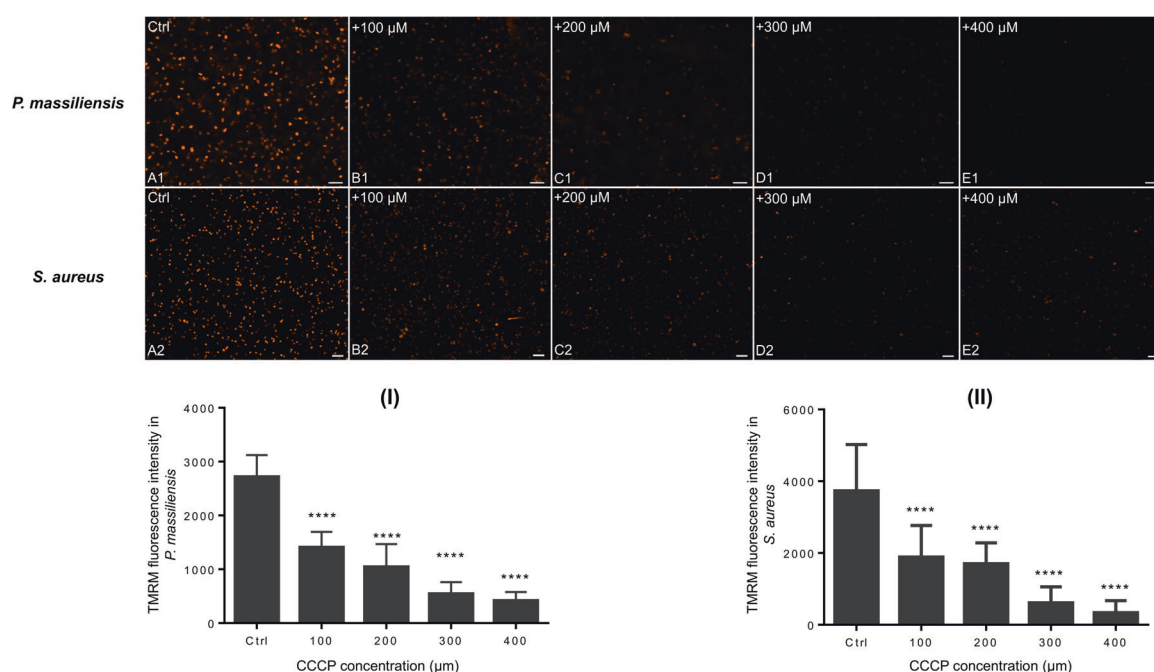
#### The depolarizing agent CCCP abolishes the potential membrane of *P. massiliensis*

*P. massiliensis* particles were incubated with a range of CCCP concentrations: 100, 200, 300, and 400  $\mu$ M. Analysis of the *S. aureus* positive control revealed the fluorescent signal generated by TMRM decreased significantly ( $p < 0.05$ ) in the presence of all concentrations of CCCP compared with the negative controls (viral particles without CCCP) (Fig. 3).

The titer of viral particles after incubation with CCCP did not show any significant difference compared with the negative control ( $10^{7.2}$  TCID<sub>50</sub>/ml). Delta-Ct between the negative control (CCCP untreated samples) and viral particles preincubated with CCCP at 400  $\mu$ M was 1.85 and 2.57 for the H0 and H3 postinfection time points, respectively. Immunofluorescence revealed a smaller number of labeled viral particles preincubated with CCCP at

400  $\mu$ M on amoeba cells (1313 and 1613 at H0 and H3, respectively) than the negative control (1658 and 1889 at H0 and H3, respectively), but the difference was not significant (Supplementary Fig. 5). We showed here that the intensity of the membrane potential is abolished following treatment with CCCP, a decoupling agent, confirming the accuracy of the existence of this electrochemical gradient.

We could only observe a lower number of viral particles on amoeba cells infected with virions preincubated with CCCP at H0 and H3 postinfection than in negative controls. This finding suggested that the membrane voltage may be involved in the infection process of amoeba cell, particularly in the early stages of infection. The membrane potential of virions may facilitate the viral DNA release into amoeba host cells. The impact of the osmolarity was previously demonstrated for *Chlorella* viruses for which inhibition of K<sup>+</sup> channels prevent the viral DNA release [36] and for bacteriophages for which genome release was suppressed by high external osmotic pressure [37]. Moreover, genes encoding photosynthetic electron transport proteins identified in phages from the *Myoviridae* and *Podoviridae* families, that infect *Prochlorococcus*, a marine cyanobacteria, may help maintaining photosynthesis during viral infection. The phage replication cycle depending on photosynthetic activity, such genes may increase the phage fitness [38]. Furthermore, a metagenomics study carried out on hydrothermal vent plumes and associated deep ocean waters demonstrated that several dsDNA viruses that infect chemolithoautotroph sulfur-oxidizing bacteria, harbor auxiliary metabolic genes for subunits of the reverse dissimilatory sulfite reductase. This enzyme oxidizes elemental sulfur, the most abundant source of chemosynthetic energy in the hydrothermal plumes and the central intermediate in the oxidation of other reduced forms of sulfur used by some sulfur-oxidizing bacteria [39]. These viruses may play an essential role in the sulfur cycle in the deep oceans and more widely, in the evolutionary dynamics of a hub in the planetary cycling of sulfur [40]. A host-derived ammonium transporter identified in a phytoplankton virus, which infects *Ostreococcus tauri* green alga, was expressed during viral infection and takes up sources of nitrogen, impacting marine nutrient cycles [41]. *Nitrosopumilus* spindle-shaped viruses, a potential important group of marine viruses that infect ammonia-oxidizing archaea might regulate dynamics, diversity and evolution of host communities and, thus, leading to a reduction in the rate of ammonia oxidation and nitrite reduction, with a global impact on carbon and nitrogen cycling [42]. In resources poor environments, viruses adopt a non lytic mode of replication allowing an adaptive way for survival of viruses [42]. Several other studies have shown that marine viral communities can adapt the diversity of their auxiliary metabolic genes to exploit and master their microbial hosts



**Fig. 3** Evaluation of the fluorescence intensity of TMRM after CCCP treatment. **A1–E1** Confocal imaging of TMRM staining following CCCP treatment of *P. massiliensis* particles. **A1** Control experiment using untreated *P. massiliensis* particles. **B1, E1**: Confocal imaging of *P. massiliensis* virions treated with different concentrations of CCCP. **A2–E2** Confocal imaging of TMRM staining after CCCP treatment of the positive control (*S. aureus*). **A2** Control experiment showing untreated *S. aureus*. **B2, E2** *S. aureus* treatment with a different concentration of CCCP. **I** Estimation of TMRM fluorescence intensity of *P. massiliensis* particles following CCCP treatment. **II** Estimation of TMRM fluorescence intensity of *S. aureus* following CCCP treatment. The scales bar correspond to 5  $\mu\text{m}$ .

according to environmental changes, to ensure keeping on an efficient viral replication in nutrient deprived conditions, especially by direct auxiliary metabolic genes-driven metabolic reprogramming [43–46]. Such genes include some involved in the central carbon metabolism, nitrogen and nutrient cyclings, soil organic matter degradation, and consequently play a key role in the global ecosystem [47–51].

#### The membrane potential of mature virions is enhanced by low concentrations and decreased by high concentrations of acetyl-CoA

In comparison to negative controls (untreated Pandoravirus particles) (Supplementary Fig. 6.A1), the TMRM fluorescent signal significantly increased in the presence of low concentrations of acetyl-CoA (0.01 mM) (Supplementary Fig. 6.F1.I) ( $p < 0.05$ ) and significantly decreased in the presence of high concentrations of acetyl CoA (0.8 mM, 0.4 mM) (Supplementary Fig. 6.B1.C1.I) ( $p < 0.05$ ). In positive control experiments using *S. aureus*, the TMRM fluorescent signal increased significantly at low concentrations of acetyl-CoA (0.1 mM) (Supplementary Fig. 6.E2.II) ( $p < 0.05$ ) and decreased significantly in the presence of high concentrations of acetyl CoA (0.8 mM, 0.4 mM) (supplementary Fig. 6.B2.C2.II) ( $p < 0.05$ ). These results demonstrate that the virion membrane potential is inherently modified following addition of variable concentrations of acetyl-CoA, a known regulator of cellular TCA cycle, and further highlighting the possible existence of a TCA cycle in *P. massiliensis*.

#### Bioinformatics analyses reveal low similarities for 8 ORFs of *P. massiliensis* with the TCA cycle

Detailed bioinformatics analyses are included in Supplementary File 2. Using DELTA-BLAST analyses against the CDD [21], low amino acid sequence similarity with some enzymes involved in cellular TCA cycle were found. Before concluding that this similarity was not significant, we searched for other predicted *P. massiliensis* gene products with similarities to other enzymes of the TCA cycle. Low similarities were found for 8 *P. massiliensis* predicted gene products:

citrate synthase (ORF577), aconitase (ORF1245), isocitrate/isopropylmalate dehydrogenase (ORF132 and ORF864), alpha-ketoglutarate decarboxylase (ORF762), (which converts alpha-ketoglutarate to succinic semialdehyde, which is an alternative pathway with the succinic semialdehyde dehydrogenase of the two steps of alpha-ketoglutarate dehydrogenase and succinate thiokinase), the succinate dehydrogenase (ORF181) and fumarase (ORF206). In addition, acetyl-CoA synthetase (ORF595), the immediate step upstream of the first step of the TCA cycle, was identified. Searching for domain similarity or structure of the predicted eight gene products with HHPRED, hmm search and Phyre2 was inconclusive. A BLASTp analysis of these 8 *P. massiliensis* ORFs revealed three of them to be conserved in some other pandoraviruses but not in all (Supplementary File 2).

One may suppose that if the *P. massiliensis* TCA genes are important for this virus, they should be conserved in other pandoraviruses too. However, it should also be noted that it is not always the case in the nature, as has been demonstrated in cyanobacteria. The TCA cycle of cyanobacteria is characterized by the absence of 2-oxoglutarate dehydrogenase, which is replaced by 2 alternative enzymes (alpha-ketoglutarate decarboxylase and succinic semialdehyde dehydrogenase). This may also be the case for *P. massiliensis*. Interestingly, not all cyanobacteria have these two alternative enzymes [52]. One can hypothesize an analogous scenario for pandoraviruses. For example, variant TCA cycles have been associated with low oxygen conditions for cyanobacteria, and reveal an essential adaptation and important plasticity to cope with readjustment to stressing environmental conditions. It also raises additional questions about a potential evolutionary advantage for the viruses harboring these genes whereas other could “survive” without them.

The search for predicted *P. massiliensis* proteins potentially involved in energy metabolism was unsuccessful for all enzymes except those in the TCA cycle, with low percentages of identity.

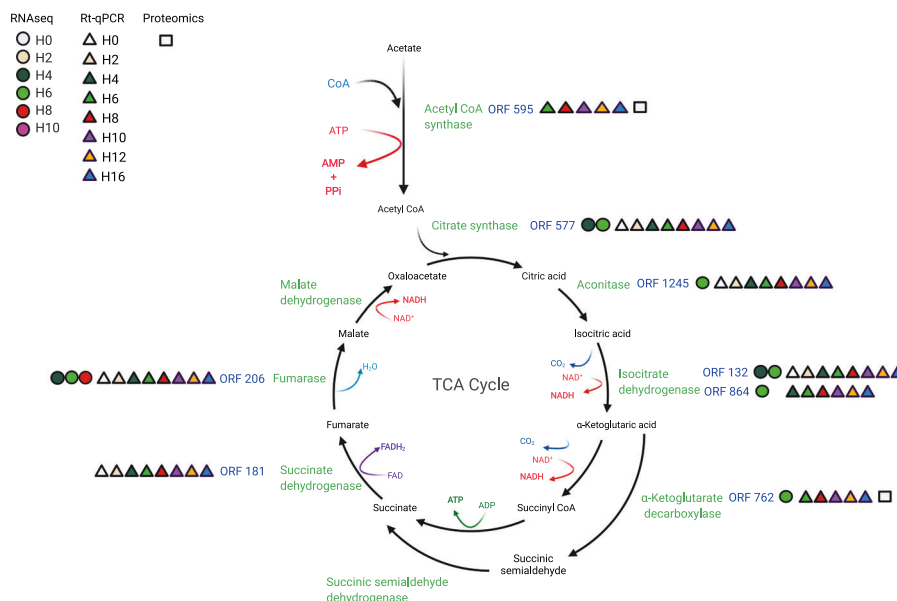
The TCA cycle is the central metabolic hub of cells. It is an exergonic catabolic energy acquisition pathway, which results in the

oxidation of an acetyl group (derived from carbon compounds) to two molecules of carbon dioxide with the concomitant harvesting of high-energy electrons. Those electrons generate a proton gradient across the inner mitochondrial membrane through oxidative phosphorylation with the aim of producing ATP catalyzed by ATP synthase [53]. The TCA cycle also provides, among other things, oxaloacetate for gluconeogenesis and metabolic intermediates for amino acid biosynthesis, nucleotide bases, cholesterol and porphyrins [53, 54]. Of note, not all eukaryotes have mitochondria [55, 56] and the TCA cycle can occur in anaerobic organisms utilizing fumarate, nitrate, or various other compounds as terminal electron acceptors instead of  $O_2$  [57, 58].

As *P. massiliensis* is neither a eubacterium nor a eukaryote, the role of a partial TCA cycle and the existence of a membrane potential in mature viral particles is currently enigmatic. In mitochondria and bacteria, the membrane potential allows cells to function as a battery and generate energy. In eukaryotic cells, mitochondrial membrane potential results in the production of ATP via the TCA cycle. For *P. massiliensis*, we could not detect the production of ATP in mature particles (data not shown). Currently, given the huge number of encoded hypothetical proteins and poor of knowledge about these recently discovered pandoraviruses, it is highly probable that these eight predicted *P. massiliensis* ORFs may be involved in alternative metabolic pathways. To date, only a few viral genes have been suggested to be involved in metabolic pathways such as in fermentation, sphingolipid biosynthesis and nitrogen metabolism [13]. *Ostreococcus tauri* viruses (Otv1-4) encode an ammonium transporter, which enables host growth rescue when cultured with ammonium as the sole nitrogen source [41]. TetV-1 encodes a mannitol metabolism enzyme, a saccharide degradation enzyme as well as other key fermentation genes [59]. In all these previous cases, the viral genes appear to be host-derived and are considered to be involved in viral manipulation of host metabolism. It has also been shown that some giant virus genomes, including pandoraviruses, contain cytochrome P450 genes which encode monooxygenase enzymes that are known to be essential in the metabolism of endogenous regulatory molecules and xenobiotics. However ancillary enzymatic redox partners encoded in these viral genomes have not been identified to-date suggesting they may be recruited from host [12]. Recently, a deep analysis of 501

environmental metagenome-assembled genomes of NCLDV revealed a diversity of metabolic genes involved in nutrient uptake, light harvesting, nitrogen metabolism, glycolysis and the TCA cycle [60]. Interestingly, BLASTp analysis of the 8 *P. massiliensis* TCA ORFs against these metagenomes [60], and against those reported by Schulz et al. [61] did not show any significant matches. Moreover, tupanviruses possess a gene encoding a putative citrate synthase [14], the first enzyme in the TCA cycle, for which no homologs have been found in any other known virus. Phylogenetic analyses indicate an independent origin of this gene in tupanviruses, suggesting it may have been acquired by tupanviruses via horizontal gene transfer from sympatric bacteria.

Phylogenetically, it has been proposed that the origin of the TCA cycle may be the reductive TCA (rTCA, also called reverse TCA) cycle an endergonic anabolic pathway [62] used by many anaerobic, microaerobic but also by some aerobic bacteria to produce carbon compounds [62–70], and is known as one of the most ancient autotrophic metabolic pathways utilizing reductive acetyl-CoA [71]. The rTCA cycle shares the same enzymes of the TCA cycle in anaerobes except for the citrate cleaving enzymes (ATP citrate lyase, or its alternatives citryl-CoA synthetase and citryl-CoA lyase). One structural study suggested that an ancestral tetrameric citryl-CoA lyase, an enzyme from the reverse TCA cycle may have evolved towards the current ATP citrate synthase [72]. Congruently, a phylometabolic analysis showed the earlier occurrence of rTCA cycle than TCA cycle [73]. Inversely, other phylogenetic analyses of the citrate cleaving enzymes reveal that they have undergone major horizontal gene transfers, suggesting that the origin of life may not be autotrophic and thus that the rTCA cycle may not be the ancestor of the TCA cycle. The rTCA cycle may have evolved as a bacterial idiosyncratic pathway [71]. However, the evolutionary history of the TCA cycle is uncertain, although it has been suggested that prior to endosymbiotic events, this pathway only operated as isolated steps [74]. Consequently, the origin of its enzymes might be associated with lateral gene transfer or duplication events, suggesting other possible metabolic functionalities than those that are currently known [75]. A selection of 1408 representative hits from cellular organisms resulting from a BLASTp analysis of the COG0473 (isocitrate/isopropylmalate dehydrogenase that was aligned with the ORF132 of *P. massiliensis*) against the nr database, allowed



**Fig. 4 Recapitulative diagram of results obtained for each step of the TCA cycle.** Schematic representation of the TCA cycle showing the predicted ORFs of *P. massiliensis* with similarities to TCA cycle enzymes, and a summary of the results provided by qRT-PCR, RNA sequencing and proteomics.



building a phylogenetic tree (Supplementary Files 3, 4, 5). This reconstruction showed that the three orthologs found in pandoraviruses (ORF132 from *P. massiliensis*, ORF339 from *P. braziliensis*, and ORF902 from *P. neocaledonia*) were closely related and grouped apart from another cluster composed of *Thermoprotei archaeon* (RLG74176.1), *Methanobacterium aggregans* (WP209583414.1), *Methanothermobacter sp* (HHY00088.1) and *Methanobacteriaceae archaeon* (NYB28041.1). However, for the other 7 ORFs of *P. massiliensis*, the absence of any robust sequence identity with known proteins, prevented phylogenetic analyses and thus the determination of the evolutionary origin of these genes. This finding again raises the question of the role of such huge genomic information reservoirs harbored by pandoraviruses, in evolutionary processes (and convergent evolution) and their role(s) in lateral gene transfers. Our data argue that new gene candidates must be considered in the search for the origin of this cycle.

### The eight ORFs of *P. massiliensis* with similarity with the TCA cycle are transcribed during the replication cycle

Transcriptomic analysis of *P. massiliensis* during the viral replication cycle in *A. castellanii* has been previously reported [15]. Briefly, a time-course of transcriptomic analysis was carried out at 30 min, 2 h, 4 h, 6 h, and 8 h post-inoculation. The data revealed that 13% of viral transcripts were detected at H4 postinfection, more than 69% at H6 postinfection and 18% detected at H8 postinfection. The total number of the viral transcriptomic reads matched with 359 *P. massiliensis* ORFs (25% of the total gene content) [15].

RNA sequencing revealed that 6 of the 8 predicted *P. massiliensis* ORFs were transcribed at different time points of the viral cycle, especially between H4 and H8 postinfection (Fig. 4). No transcripts were detected for ORF595 (putative acetyl-coenzyme A synthetase) or ORF181 (putative succinate dehydrogenase). qRT-PCR revealed that the 8 ORFs were transcribed at different time points of the viral cycle. For ORFs 595, 577, 1245, 132, 864, 762, and 206, the lowest Ct values were globally found between H8 and H16 postinfection. (Supplementary File 6).

Finally, transcriptomics analyses demonstrated that the predicted *P. massiliensis* TCA ORFs were all transcribed at the same time points, especially at the end of the developmental cycle of the virus (Fig. 4).

### Two of eight ORFs of *P. massiliensis* with similarity with the TCA cycle are expressed in mature virions

Proteomic analysis allowed us to identify 623 proteins with significant matches to known proteins in the amoeba/virus database. Among them, 96.8% (603 proteins) were derived from the host amoeba organism and only 3.2% (20 proteins) were identified as viral proteins. In a previous work, 162 protein sequences have been identified in mature virions [15]. Two *P. massiliensis* proteins of the eight predicted to be potentially involved in a partial TCA pathway were identified by this analysis in mature particles: ORF762 (putative  $\alpha$ -ketoglutarate decarboxylase) and ORF595 (putative acetyl-coenzyme A synthetase) with an identity percentage of 100% (Fig. 4).

### Characterization of enzymatic functions

Genes encoding the predicted enzymes of interest were synthesized, cloned and transformed into competent *E. coli* for recombinant protein expression. Soluble proteins were obtained for ORFs 577 (citrate synthase), 1245 (aconitate hydratase), 132, 864 (isocitrate dehydrogenase), 787 and 1146 ( $\alpha$ -ketoglutarate decarboxylases). However, no enzymatic activity could be observed for ORFs 577, 1245, 787 and 1146 (data not shown). We investigated the potential isocitrate dehydrogenase activity of ORFs 132 and 864. We determined a specific activity of 4 mU/mg for ORF132 (Supplementary Fig. 7), but no activity for ORF864. Kinetic assays were also performed to evaluate the catalytic

parameters of ORF132. According to Michaelis–Menten equation fitting ( $R^2 = 0.993$ ), the following parameters were determined:  $k_{cat} = 6.8 \times 10^{-4} \text{ s}^{-1}$ ,  $K_m = 51.8 \mu\text{M}$  and  $k_{cat}/K_m = 13.12 \text{ s}^{-1} \text{ M}^{-1}$  (Supplementary Fig. 8). The IDH activity assessed on purified mature virions was similar to those of the recombinant ORF132. In parallel, we performed the same analysis of human IDH from Sigma-Aldrich (St. Louis, MS, USA) and could determine a specific activity of 6.3 U/mg using the previously mentioned kit, as well as a  $k_{cat} = 16.3 \text{ s}^{-1}$ ,  $K_m = 585.4 \mu\text{M}$  and  $k_{cat}/K_m = 2.78 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$  with an  $R^2$  of 0.997. In the TCA cycle, isocitrate dehydrogenase converts the isocitrate to  $\alpha$ -ketoglutarate in the presence of the  $\text{NAD}^+$  or  $\text{NADP}^+$  cofactor. In nature, isocitrate dehydrogenase catalyzes a catabolic reaction, during which  $\text{NAD}^+$  abstracts a hydride ion, a highly stereospecific enzymatic mechanism. Thus, the functionality of *P. massiliensis* isocitrate dehydrogenase cannot simply be the result of chance.

In cells, the IDH step of the TCA cycle is an irreversible reaction, with an overall estimated free energy of  $-8.4 \text{ kJ/mol}$  [76]. It is regulated by substrate availability, product inhibition, and competitive feedback inhibition by ATP [77]. Isocitrate binds within the enzyme active site, involving 8 key amino acids. The metal ion  $\text{Mg}^{2+}$  or  $\text{Mn}^{2+}$  binds to three conserved Arg residues through hydrogen bond networks. The cofactor  $\text{NAD}^+$  or  $\text{NADP}^+$  binds within four regions with similar properties among the IDH enzymes, located around amino acids [250–260], [280–290], [300–330], and [365–380] [78]. ORF132 is 146 amino acids long, while the known IDH from *E. coli* is 417 long (QJ224410.1). IDH is typically multimeric [79] suggesting that the pandoraviral enzyme may also form multimers. If the pandoraviral IDH is an ancestral form of the current eukaryotic IDH, then it is not surprising that the  $K_m$  and  $K_{cat}$  of ORF132 are low. Evolutionary, enzymes involved in the TCA cycle have evolved towards optimal performance with higher  $K_m$  and  $K_{cat}$  values.

### CONCLUSION AND PERSPECTIVES

Experimental evidence of a virus encoding enzymes of a partial TCA cycle opens a new horizon in giant virus research. For each experiment, critical negative controls were carried out to avoid false positive results. Membrane potential and in particular, membrane depolarization, have been previously described in *Chlorella* viruses, and have been linked to  $\text{K}^+$  protein channels [80, 81]. Therein, the fusion of the virus internal membrane with the host plasma membrane results in the membrane depolarization that occurs early in *Chlorella* virus-host interactions. This phenomenon, lowering the energy barrier, allows the viral DNA release into the host. However, it appears to differ from those observed in pandoraviruses. Indeed, pandoraviruses seem to harbor predicted genes with low similarity with all enzymes of the cellular TCA cycle and most importantly, a functional isocitrate dehydrogenase, a key enzyme of this cycle. Moreover, the membrane potential of *P. massiliensis* virions is affected by acetyl-CoA, a well-known regulator of the TCA cycle. One may assume that a complete TCA cycle can occur if all the enzymatic components are present in a system. Herein, our findings suggest the possible existence of an autonomous, at least partial, TCA pathway potentially involved in energy production, or in another, unknown metabolic pathway, in a virus. As the TCA cycle plays a key role in anabolism and catabolism of many biomolecules, it also can also be hypothesized that these viral genes may be involved in redirecting amoeba carbon metabolism.

It is known that viruses are major drivers in nutrient and energy recycling within natural environments and in the ecology and evolution of their cellular hosts [82–85]. In the setting of host-virus interactions, the ability of the viruses to generate energy would increase their own fitness. Indeed, one can imagine that the delivery of reducing power ( $\text{NADH}$  or  $\text{NADPH}$ ) through the

putative virus-encoded TCA enzymes into the host cell upon infection can provide a significant increase of direct usable energy to sustain the initial step of viral replication without affecting host cell needs. This can be comparable to cyanoviruses, which carry horizontally transferred auxiliary metabolic genes that presumably maintain functional photosynthetic machinery during infection and supply energy to the bacterial host, as transcription of host genes drops during infection [38, 86–88]. The prevalence of such auxiliary metabolic genes in the environment [45, 89, 90] raises important questions regarding their contribution in the biosphere [45]. In the light of this study, we have to rethink and take into account viral input(s) in energy storage and flux pathways within the microbial food web. Future experiments are clearly needed for understanding fully the existence of a TCA cycle in such viruses and are underway in our laboratories. It should be noted the limit of the standard comparative genomics tool in identifying novel, or small and very phylogenetically divergent sequences. Experiments investigating the crystal structures of the viral enzymes, especially those of IDH, will be of particular interest for advancing comprehension of this mysterious Pandoravirus. Moreover, the use of yeast and bacterial complementation experiments to confirm the predicted enzymatic functions seems to be the next logical steps. The database of newly discovered giant viruses is constantly growing and constitutes an important resource for the search for similar predicted enzymatic functions potentially involved in metabolic pathways.

In conclusion, *P. massiliensis* undermines the last known historical viral hallmark, the lack of the Lipman system. Thus, our findings presented herein raise critical questions concerning whether pandoraviruses can still be accurately classified biologically as viruses and renew deep-rooted arguments regarding the living nature of viruses in general.

## REFERENCES

- La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, et al. A giant virus in amoebae. *Science*. 2003;299:2033.
- Rolland C, Andreani J, Louazani AC, Aherfi S, Francis R, Rodrigues R, et al. Discovery and further studies on giant viruses at the IHU Mediterranean infection that modified the perception of the virosphere. *Viruses*. 2019;11:312.
- Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of Mimivirus. *Science*. 2004;306:1344–50.
- Philippe N, Legendre M, Dautre G, Coute Y, Poirot O, Lescot M, et al. *Pandoraviruses*: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science*. 2013;341:281–6.
- Brahim Belhauari D, Baudoin JP, Gnankou F, Di Pinto F, Colson P, Aherfi S, et al. Evidence of a cellulosic layer in pandoravirus massiliensis tegument and the mystery of the genetic support of its biosynthesis. *Front Microbiol*. 2019;10:2932.
- La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, et al. The virophage as a unique parasite of the giant mimivirus. *Nature*. 2008;455:100–4.
- Levasseur A, Bekliz M, Chabriere E, Pontarotti P, La Scola B, Raoult D. MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature*. 2016;531:249–52.
- Bekliz M, Azza S, Seligmann H, Decloquement P, Raoult D, La Scola B. Experimental analysis of mimivirus translation initiation factor 4a reveals its importance in viral protein translation during infection of acanthamoeba polyphaga. *J Virol*. 2018;92:e00337–18.
- Abraham J, Silva L, Silva LS, Bou Khalil JY, Rodrigues R, Arantes T, et al. Tailed giant *Tupanvirus* possesses the most complete translational apparatus of the known virosphere. *Nat Commun*. 2018;9:749.
- Schulz F, Yutin N, Ivanova NN, Ortega DR, Lee TK, Vierheilig J, et al. Giant viruses with an expanded complement of translation system components. *Science*. 2017;356:82–5.
- Lamb DC, Lei L, Warrilow AG, Lepesheva GI, Mullins JG, Waterman MR, et al. The first virally encoded cytochrome p450. *J Virol*. 2009;83:8266–9.
- Lamb DC, Follmer AH, Goldstone JV, Nelson DR, Warrilow AG, Price CL, et al. On the occurrence of cytochrome P450 in viruses. *Proc Natl Acad Sci USA*. 2019;116:12343–52.
- Vardi A, Haramaty L, Van Mooy BA, Fredricks HF, Kimmance SA, Larsen A, et al. Host-virus dynamics and subcellular controls of cell fate in a natural coccolithophore population. *Proc Natl Acad Sci USA*. 2012;109:19327–32.
- Araújo Lima Rodrigues R, Arantes TS, Oliveira GP, Dos Santos Silva LK, Abraham JS. The Complex Nature of *Tupanviruses*. *Adv Virus Res*. 2019;103:135–66.
- Aherfi S, Andreani J, Baptiste E, Oumessoum A, Dornas FP, Dos Santos Andrade AC, et al. A Large Open Pangenome and a Small Core Genome for Giant *Pandoraviruses*. *Front Microbiol*. 2018;9:1486.
- La Scola B, Rydkina L, Ndiokubwayo JB, Vene S, Raoult D. Serological differentiation of murine typhus and epidemic typhus using cross-adsorption and Western blotting. *Clin Diagn Lab Immunol*. 2000;7:612–6.
- Andreani J, Arnault JP, Bou Khalil JY, Abraham J, Tomei E, Vial E, et al. Atypical Cowpox Virus Infection in Smallpox-Vaccinated Patient, France. *Emerg Infect Dis*. 2019;25:212–9.
- Trocha LK, Stobienia O. Response of *Acanthamoeba castellanii* mitochondria to oxidative stress. *Acta Biochim Pol*. 2007;54:797–803.
- Reed LJ, Muench H. A simple method of estimating fifty per cent endpoints. *Am J Hyg*. 1938;27:493–7.
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res*. 2011;39:D225–9.
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*. 2017;45:D200–3.
- Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, et al. A completely reimplemented MPI bioinformatics Toolkit with a new HHpred server at its core. *J Mol Biol*. 2018;430:2237–43.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The PyMol web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10:845–58.
- Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinform*. 2011;12:124. <https://doi.org/10.1186/1471-2105-12-124>.
- Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28:33–6.
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinform*. 2003;4:41.
- Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform*. 2004;5:113.
- Price MN, Dehal PS, Arkin AP. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS ONE*. 2010;5:e9490.
- Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49:W293–6.
- Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics*. 2007;23:1289–91.
- Remy B, Plener L, Decloquement P, Armstrong N, Elias M, Daude D, et al. Lactonase Specificity Is Key to Quorum Quenching in *Pseudomonas aeruginosa*. *Front Microbiol*. 2020;11:762.
- Zorova LD, Popkov VA, Plotnikov EY, Silachev DN, Pevzner IB, Jankauskas SS, et al. Mitochondrial membrane potential. *Anal Biochem*. 2018;552:50–9.
- Edwards SW, Lloyd D. Properties of mitochondria isolated from cyanide-sensitive and cyanide-stimulated cultures of *Acanthamoeba castellanii*. *Biochemical J*. 1978;174:203–11.
- Leger MM, Gawryluk RM, Gray MW, Roger AJ. Evidence for a hydrogenosomal-type anaerobic ATP generation pathway in *Acanthamoeba castellanii*. *PLoS ONE*. 2013;8:e69532.
- Hug LA, Stechmann A, Roger AJ. Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. *Mol Biol Evol*. 2010;27:311–24.
- Neupärtl M, Meyer C, Woll I, Frohns F, Kang M, Van Etten JL, et al. *Chlorella* viruses evoke a rapid release of K<sup>+</sup> from host cells during the early phase of infection. *Virology*. 2008;372:340–8.
- Evilevitch A, Lavelle L, Knobler CM, Raspaud E, Gelbart WM. Osmotic pressure inhibition of DNA ejection from phage. *Proc Natl Acad Sci USA*. 2003;100:9292.
- Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc Natl Acad Sci USA*. 2004;101:11013–8.
- McCormell TM. Geochemical constraints on primary productivity in submarine hydrothermal vent plumes. *Deep Sea Res 1 Oceanogr Res Pap*. 2000;47:85–101.
- Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ. Sulfur oxidation genes in diverse deep-sea viruses. *Science*. 2014;344:757–60.
- Monier A, Chambouvet A, Milner DS, Attah V, Terrado R, Lovejoy C, et al. Host-derived viral transporter protein for nitrogen uptake in infected marine phytoplankton. *Proc Natl Acad Sci USA*. 2017;114:E7489–98.
- Kim JG, Kim SJ, Cvirkaitė-Krupovic V, Yu WJ, Gwak JH, Lopez-Perez M, et al. Spindle-shaped viruses infect marine ammonia-oxidizing thaumarchaea. *Proc Natl Acad Sci USA*. 2019;116:15645–50.



43. Coutinho FH, Silveira CB, Gregoracci GB, Thompson CC, Edwards RA, Brussaard CPD, et al. Marine viruses discovered via metagenomics shed light on viral strategies throughout the oceans. *Nat Commun*. 2017;8:15955.
44. Howard-Varona C, Lindback MM, Bastien GE, Solonenko N, Zayed AA, Jang H, et al. Phage-specific metabolic reprogramming of virocells. *ISME J*. 2020;14: 881–95.
45. Hurwitz BL, U'Ren JM. Viral metabolic reprogramming in marine ecosystems. *Curr Opin Microbiol*. 2016;31:161–8.
46. Hurwitz BL, Hallam SJ, Sullivan MB. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol*. 2013;14:R123.
47. Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol*. 2007;5:801–12.
48. Zimmerman AE, Howard-Varona C, Needham DM, John SG, Worden AZ, Sullivan MB, et al. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat Rev Microbiol*. 2020;18:21–34.
49. Brum JR, Sullivan MB. Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol*. 2015;13:147–59.
50. Trubl G, Jang HB, Roux S, Emerson JB, Solonenko N, Vik DR, et al. Soil viruses are underexplored players in ecosystem carbon processing. *mSystems*. 2018;3: e00076–18.
51. Crummett LT, Puxty RJ, Weihe C, Marston MF, Martiny JBH. The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. *Virology*. 2016;499:219–29.
52. Steinhauser D, Fernie AR, Araújo WL. Unusual cyanobacterial TCA cycles: not broken just different. *Trends Plant Sci*. 2012;17:503–9.
53. Berg JM, Tymoczko JL, Stryer L. *Biochemistry*. 5th ed. New York: W H Freeman; 2002. <https://www.ncbi.nlm.nih.gov/books/NBK21154/>.
54. Krebs HA, Johnson WA. Acetopyruvic acid (alpha-gamma-diketovaleric acid) as an intermediate metabolite in animal tissues. *Biochem J*. 1937;31:772–9.
55. Muller M. Energy metabolism of protozoa without mitochondria. *Annu Rev Microbiol*. 1988;42:465–88.
56. Muller M. The hydrogenosome. *J Gen Microbiol*. 1993;139:2879–89.
57. Tielens AG, Van, Hellemond JJ. The electron transport chain in anaerobically functioning eukaryotes. *Biochim Biophys Acta*. 1998;1365:71–8.
58. Hedderich R, Klimmek O, Kröger A, Dirmeier R, Keller M, Stetter KO. Anaerobic respiration with elemental sulfur and with disulfides. *FEMS Microbiol Rev*. 1999; 22:353–81.
59. Schvarcz CR, Steward GF. A giant virus infecting green algae encodes key fermentation genes. *Virology*. 2018;518:423–33.
60. Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat Commun*. 2020;11:1710.
61. Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh DA, Denef VJ, et al. Giant virus diversity and host interactions through global metagenomics. *Nature*. 2020;578:432–6.
62. Aoshima M. Novel enzyme reactions related to the tricarboxylic acid cycle: phylogenetic/functional implications and biotechnological applications. *Appl Microbiol Biotechnol*. 2007;75:249–55.
63. Evans MC, Buchanan BB, Arnon DI. A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. *Proc Natl Acad Sci USA*. 1966;55:928–34.
64. Buchanan BB, Arnon DI. A reverse KREBS cycle in photosynthesis: consensus at last. *Photosynth Res*. 1990;24:47–53.
65. Shiba H, Kawasumi T, Igarashi Y, Kodama T, Minoda Y. The CO<sub>2</sub> assimilation via the reductive tricarboxylic acid cycle in an obligately autotrophic, aerobic hydrogen-oxidizing bacterium, *Hydrogenobacter thermophilus*. *Arch Microbiol*. 1985;141:198–203.
66. Hugler M, Huber H, Stetter KO, Fuchs G. Autotrophic CO<sub>2</sub> fixation pathways in archaea (*Crenarchaeota*). *Arch Microbiol*. 2003;179:160–73.
67. Hugler M, Huber H, Molyneux SJ, Vetriani C, Sievert SM. Autotrophic CO<sub>2</sub> fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum *Aquificae*: evidence for two ways of citrate cleavage. *Environ Microbiol*. 2007;9:81–92.
68. Williams TJ, Zhang CL, Scott JH, Bazylinski DA. Evidence for autotrophy via the reverse tricarboxylic acid cycle in the marine magnetotactic coccus strain MC-1. *Appl Environ Microbiol*. 2006;72:1322–9.
69. Bayer B, Saito MA, McIlvin MR, Lucker S, Moran DM, Lankiewicz TS, et al. Metabolic versatility of the nitrite-oxidizing bacterium *Nitrospira marina* and its proteomic response to oxygen-limited conditions. *ISME J*. 2021;15:1025–39.
70. Luecker S, Nowka B, Rattei T, Spieck E, Daims H. The genome of *Nitrospina gracilis* illuminates the metabolism and evolution of the major marine nitrite oxidizer. *Front Microbiol*. 2013;4:27.
71. Becerra A, Rivas M, Garcia-Ferris C, Lazcano A, Pereto J. A phylogenetic approach to the early evolution of autotrophy: the case of the reverse TCA and the reductive acetyl-CoA pathways. *Int Microbiol*. 2014;17:91–7.
72. Verschueren KHG, Blanchet C, Felix J, Dansercoer A, De Vos D, Bloch Y, et al. Structure of ATP citrate lyase and the origin of citrate synthase in the Krebs cycle. *Nature*. 2019;568:571–5.
73. Braakman R, Smith E. The emergence and early evolution of biological carbon-fixation. *PLoS Comput Biol*. 2012;8:e1002455.
74. Schnarrenberger C, Martin W. Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. *Eur J Biochem*. 2002;269:868–83.
75. Cavalcanti JH, Esteves-Ferreira AA, Quinhones CG, Pereira-Lima IA, Nunes-Nesi A, Fernie AR, et al. Evolution and functional implications of the tricarboxylic acid cycle as revealed by phylogenetic analysis. *Genome Biol Evol*. 2014;6:2830–48.
76. Garrett RH. *Grisham CM biochemistry*. 5th ed. Boston, MA: Cengage learning; 2012.
77. Maeting I, Schmidt G, Sahm H, Stahmann KP. Role of a peroxisomal NADP-specific isocitrate dehydrogenase in the metabolism of the riboflavin overproducer *Ashbya gossypii*. *J Mol Catal B Enzym*. 2000;410:335–43.
78. Fedoy AE, Yang N, Martinez A, Leiros HK, Steen IH. Structural and functional properties of isocitrate dehydrogenase from the psychrophilic bacterium *Desulfotalea psychrophila* reveal a cold-active enzyme with an unusual high thermal stability. *J Mol Biol*. 2007;372:130–49.
79. Steen IH, Lien T, Birkeland NK. Biochemical and phylogenetic characterization of isocitrate dehydrogenase from a hyperthermophilic archaeon, *Archaeoglobus fulgidus*. *Arch Microbiol*. 1997;168:412–20.
80. Frohns F, Kasman A, Kramer D, Schafer B, Mehmel M, Kang M, et al. Potassium ion channels of *Chlorella* viruses cause rapid depolarization of host cells during infection. *J Virol*. 2006;80:2437–44.
81. Plugge B, Gazzarini S, Nelson M, Cerana R, Van Etten JL, Derst C, et al. A potassium channel protein encoded by *Chlorella* virus PBCV-1. *Science*. 2000; 287:1641–4.
82. Forterre P. The virocell concept and environmental microbiology. *ISME J*. 2013;7:233–6.
83. Hambly E, Suttle CA. The virosphere, diversity, and genetic exchange within phage communities. *Curr Opin Microbiol*. 2005;8:444–50.
84. Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol*. 2007;5:801–12.
85. Kuzakov Y, Mason-Jones K. Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biol Biochem*. 2018;127:305–17.
86. Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature*. 2005;438:86–9.
87. Sullivan MB, Huang KH, Ignacio-Espinoza JC, Berlin AM, Kelly L, Weigele PR, et al. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ Microbiol*. 2010;12: 3035–56.
88. Zeidner G, Bielawski JP, Shmoish M, Scanlan DJ, Sabehi G, Bèjà O. Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. *Environ Microbiol*. 2005;7:1505–13.
89. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, et al. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci USA*. 2011;108:E757.
90. Puxty RJ, Millard AD, Evans DJ, Scanlan DJ. Shedding new light on viral photosynthesis. *Photosynth Res*. 2015;126:71–97.

## ACKNOWLEDGEMENTS

We sincerely thank Christelle Desnues for her ideas, and helpful advises on the discussion of the article. We thank Rita Jaafar for her assistance in editing and formatting phylogenetic trees.

## AUTHOR CONTRIBUTIONS

BLS conceived the project. SA, DBB, LP, JPB, and PD performed the experiments. PC, AL, EC, DR, and BLS supervised the study. SA, DBB, LP, PD, and BLS wrote the paper. JA, DCL revised and improved the paper.

## FUNDING

This work was supported by the French Government under the “Investments for the Future” program managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03. It was also supported by Région Provence-Alpes-Côte d’Azur and European funding FEDER PRIMMI (Fonds Européen de Développement Régional—Plateformes de Recherche et d’Innovation Mutualisées Méditerranée Infection).

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41396-021-01117-3>.

**Correspondence** and requests for materials should be addressed to Bernard La Scola.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021