Vol 455|4 September 2008|doi:10.1038/nature07218

LETTERS

The virophage as a unique parasite of the giant mimivirus

Bernard La Scola[1]*, Christelle Desnues[1]*, Isabelle Pagnier[1], Catherine Robert[1], Lina Barrassi[1], Ghislain Fournous[1], Michèle Merchat[2], Marie Suzan-Monti[1], Patrick Forterre[3,4], Eugene Koonin[5] & Didier Raoult[1]

Viruses are obligate parasites of Eukarya, Archaea and Bacteria. *Acanthamoeba polyphaga* mimivirus (APMV) is the largest known virus; it grows only in amoeba and is visible under the optical microscope. Mimivirus possesses a 1,185-kilobase double-stranded linear chromosome whose coding capacity is greater than that of numerous bacteria and archaea[1–3]. Here we describe an icosahedral small virus, Sputnik, 50 nm in size, found associated with a new strain of APMV. Sputnik cannot multiply in *Acanthamoeba castellanii* but grows rapidly, after an eclipse phase, in the giant virus factory found in amoebae co-infected with APMV[4]. Sputnik growth is deleterious to APMV and results in the production of abortive forms and abnormal capsid assembly of the host virus. The Sputnik genome is an 18.343-kilobase circular double-stranded DNA and contains genes that are linked to viruses infecting each of the three domains of life Eukarya, Archaea and Bacteria. Of the 21 predicted protein-coding genes, eight encode proteins with detectable homologues, including three proteins apparently derived from APMV, a homologue of an archaeal virus integrase, a predicted primase–helicase, a packaging ATPase with homologues in bacteriophages and eukaryotic viruses, a distant homologue of bacterial insertion sequence transposase DNA-binding subunit, and a Zn-ribbon protein. The closest homologues of the last four of these proteins were detected in the Global Ocean Survey environmental data set[5], suggesting that Sputnik represents a currently unknown family of viruses. Considering its functional analogy with bacteriophages, we classify this virus as a virophage. The virophage could be a vehicle mediating lateral gene transfer between giant viruses.

The original strain of APMV, mimivirus, was obtained from a cooling tower in Bradford, UK. Its size challenged the definition of a virus[6] and led to the idea that giant viruses might be an uncharacterized but important part of the biosphere. We isolated a new strain of APMV, by inoculating *A. polyphaga* with water from a cooling tower, in Paris. We denoted this new strain mamavirus because it seemed to be even larger than mimivirus[2] when observed by transmission electron microscopy. The main features of mamavirus closely resembled those described for mimivirus, including the formation of a giant viral factory and the typical particle morphology with a multilayered membrane covered with fibrils[4]. We also observed unknown icosahedral small viral particles, 50 nm in size, in virus factories and in the cytoplasm of the infected cells (Fig. 1). Considering the association of this newly detected virus with mamavirus, we named it Sputnik.

Sputnik did not multiply when inoculated into *A. castellanii* (Supplementary Information and Supplementary Table 4). However, this virus did grow, as demonstrated by transmission electron microscopy and polymerase chain reaction, in *A. castellanii* co-infected with mimivirus or mamavirus (Supplementary Information

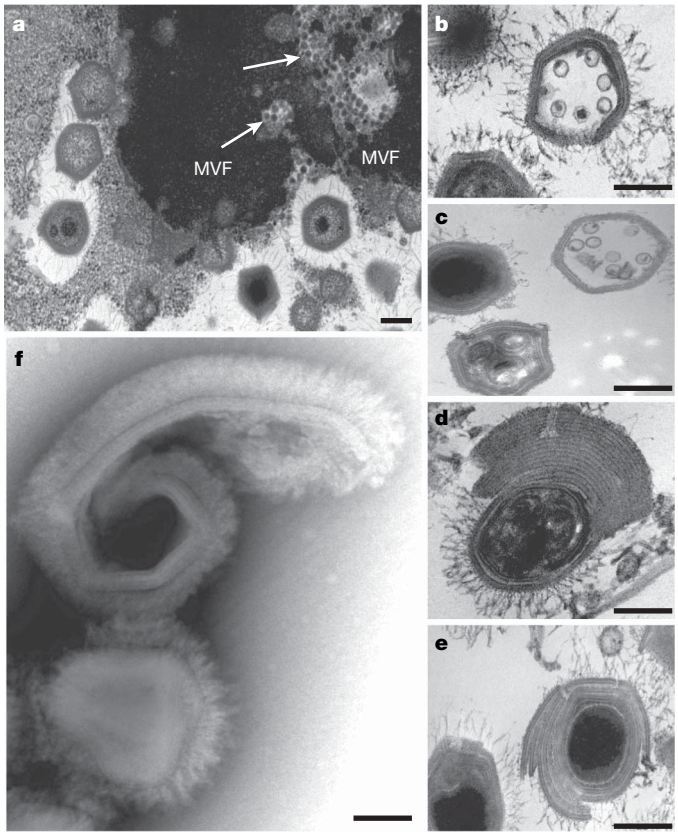

Figure 1 | Different morphological aspects of mamavirus and Sputnik. **a–e**, Observations by transmission electron microscopy; **f**, observation by negative staining electron microscopy. **a**, Mamavirus virus factory (MVF) with mamavirus particles at different stages of maturation. Clumps of Sputnik particles (arrows) are observed within MVF. **b**, In some cases, Sputnik is observed within mamavirus capsids. **c**, Defective particles are produced. **d–f**, Co-infection with mamavirus and Sputnik results in abnormal morphology of mamavirus particles, such as membrane accumulation at one side (**d**), membrane accumulation around the particles (**e**), or open particles (**f**). Scale bars, 200 nm.

[1]URMITE, Centre National de la Recherche Scientifique UMR IRD 6236, Faculté de Médecine, Université de la Méditerranée, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France. [2]Climespace, 185 Rue de Bercy, 75012 Paris, France. [3]Biologie Moléculaire du Gène chez les Extrêmophiles, Institut de Génétique et Microbiologie, Bâtiment 409, Université Paris Sud, Centre d'Orsay, 91405 Orsay Cedex, France. [4]Biologie Moléculaire du Gène chez les Extrêmophiles, Département de Microbiologie, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France. [5]National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health, Building 38A, Room 5N503, 8600 Rockville Pike, Bethesda, Maryland 20894, USA.
*These authors contributed equally to this work.

©2008 Macmillan Publishers Limited. All rights reserved

and Supplementary Table 4). Sputnik and mamavirus were produced within the same viral factory with different kinetics and at different specific locations. Sputnik was produced earlier than APMV (Fig. 2). Sputnik co-infection was associated with a significant increase in the formation of abnormal mamavirus virions, characterized by partial thickening of the capsid (11% rather than 1%, $P = 0.0029$). In the regular mamavirus virions, the capsid layer was 40 nm thick; in contrast, in the presence of Sputnik, the thickness of the capsid wall could reach 240 nm (Fig. 1). In most cases, several capsid layers accumulated asymmetrically at one pole of the viral particle. Some of these abnormal particles seemed to be mature and to harbour fibrils only on the normal part of the capsid. Only a small fraction of the mamavirus particles encapsidated Sputnik virions (Fig. 1). However, co-inoculation of mamavirus with Sputnik resulted in a roughly 70% decrease in the yield of infective mamavirus particles and a threefold decrease in amoeba lysis at 24 h. These findings showed that Sputnik is a parasite of mamavirus that substantially affects the reproduction of the host virus.

The *Acanthamoeba castellanii mamavirus* genome (C.D., B.L.S., C.R., G.F. and D.R., unpublished observations) is about 1,200 kilobase pairs in size. Its genome is highly AT-rich (A + T content ≈ 72%). Orthologues to mimivirus open reading frames (ORFs) were detected for 99% of the predicted mamavirus genes, with amino-acid identity ranging from 75% to 100%. Thus, mamavirus is closely related to mimivirus and could be considered a second strain of APMV. Sputnik has an 18,343-base-pair (bp) circular double-stranded DNA genome, with 21 predicted protein-coding genes ranging in size from 88 to 779 amino-acid residues (Table 1 and Fig. 3). The organization of the Sputnik genome is typical of viral genomes, namely a tight arrangement but little overlap of the ORFs. The high A + T content (73%) of the Sputnik genome is similar to that of APMV. Sputnik samples were resolved by two-dimensional gel electrophoresis within a pI range of 3–10 (Fig. 3). The most abundant of the detected protein spots, analysed by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) mass spectrometry, corresponded to ORF 20; ORF 08 and ORF 19 proteins were identified once each. These results were corroborated by western blot analysis with a mouse anti-serum against purified Sputnik (Supplementary Fig. 1). Thus, ORF 20 most probably encodes the major capsid protein of Sputnik, whereas ORFs 08 and 19 encode minor virion proteins.

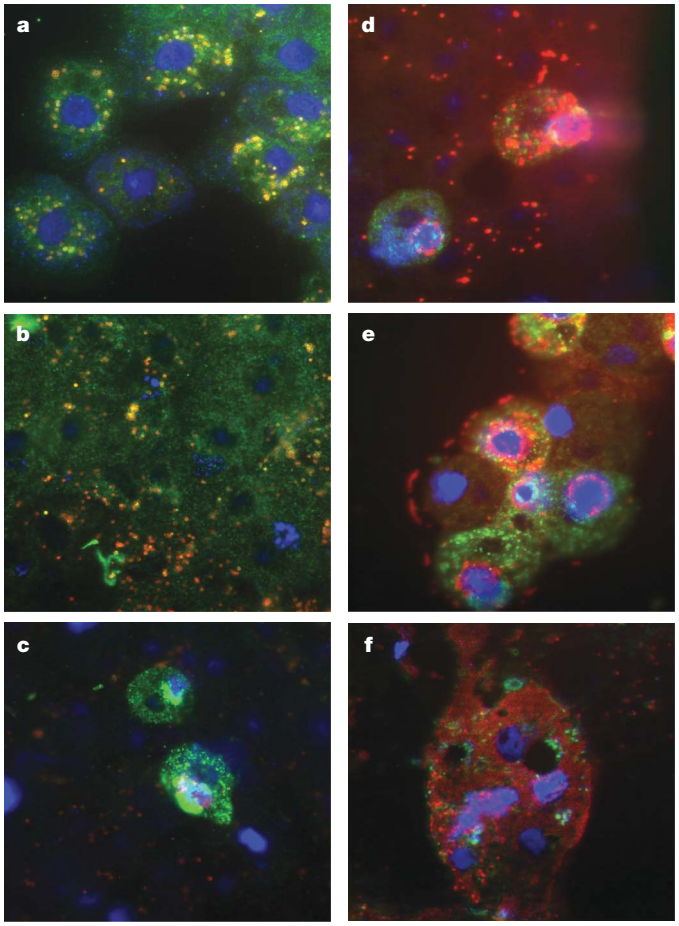

Figure 2 | Sputnik propagation in mamavirus-infected amoebae. *A. castellanii* cells were infected with a mixture of mamavirus and Sputnik. Indirect immunofluorescence labelling was performed with rabbit anti-mimivirus serum (red) and mouse anti-Sputnik serum (green), and nucleic acids were stained with 4,6-diamidino-2-phenylindole (DAPI; blue). **a**, Numerous Sputnik virions entered the cytoplasm at 30 min after infection. **b**, At 4 h after infection, the first viral factories were seen as distinct, strongly stained patches. No viral particles could be seen in these cells, indicating an eclipse phase. **c**, At 6 h after infection, the viral factories expanded and were homogenously and strongly stained with DAPI. Sputnik production was detected at one side of the viral factory, but no mamavirus virions. **d**–**f**, At 8 h after infection (**d**), mamavirus production was observed; this increased extensively at 12 h (**e**) and 16 h (**f**) after infection.

Genomes of many viruses contain a high proportion of 'ORFan' genes; that is, genes without detectable homologues in current sequence databases. The genome of Sputnik is no exception because most of its encoded proteins (13 of 21) are ORFans. The eight non-ORFan proteins have viral/plasmid, bacterial or eukaryotic homologues, and/or homologues from the environmental Global Ocean Survey (GOS) data set (Table 1). Three of the Sputnik predicted proteins (ORFs 6, 12 and 13) were most closely related to mimivirus/mamavirus gene products. The proteins encoded in ORFs 12 and 13 were equally similar to their respective homologues from the mimivirus and the mamavirus (Supplementary Table 3), whereas ORF 6 was more closely related to the mamavirus homologue. The most plausible model is therefore that Sputnik acquired a portion of the gene (or the complete gene, which was further partly eliminated) from mamavirus after its divergence from the common ancestor with mimivirus.

Specifically, ORF 12 is uncharacterized, whereas ORFs 6 and 7 encode paralogous proteins containing highly conserved collagen triple-helix motifs[7]. The protein encoded by ORF 13 consists of two domains implicated in viral DNA replication. The carboxy-terminal domain of this protein is a superfamily 3 helicase that is highly conserved and clusters with homologues from nucleocytoplasmic large DNA viruses (NCLDVs)[7] in phylogenetic trees (Fig. 3 and Supplementary Figs 2 and 3). The amino-terminal portion of ORF 13 protein is a previously unobserved domain for which homologues with high similarity were detected only among proteins from the GOS data set and which, on the basis of the presence of a signature sequence motif, could be predicted to represent a highly derived version of the archaeo-eukaryotic primase (Supplementary Fig. 4). The ORF 3 protein showed limited similarity to a packaging ATPase of the FtsK–HerA superfamily that is found in all NCLDVs and many bacteriophages[5,8,9] (Fig. 3 and Supplementary Fig. 5). ORF 14, which is adjacent to the primase–helicase gene, encodes a protein containing a Zn-ribbon motif that is significantly similar to that in several proteins in the GOS data set (Table 1 and Supplementary Fig. 6), and ORF 4 also encodes a Zn-ribbon protein without highly conserved homologues. ORF 17 encodes a protein with homologues in the GOS data set that belong to the family of bacterial insertion sequence transposase DNA-binding subunits/domains (transposase A proteins) (Table 1, Fig. 3 and Supplementary Fig. 7). Finally, ORF 10 protein showed significant sequence similarity to integrases of the tyrosine recombinase family from archaeal viruses and proviruses, a relationship that was further supported by phylogenetic analysis (Fig. 3 and Supplementary Fig. 8).

Two genes implicated in essential functions in viral genome replication and packaging (ORFs 13 and 3, respectively) and a gene with a potential role in expression regulation (ORF 14) are most closely

©2008 Macmillan Publishers Limited. All rights reserved

related to genes from the GOS data set. Given that the primase–helicase and the FtsK-like ATPase are typical viral genes, it seems likely that Sputnik is linked to an unknown family of viruses, perhaps related to NCLDVs, that is abundantly represented among the marine metagenomic sequences but not in other current sequence databases.

Thus, the Sputnik genome contains genes evolutionarily related to at least three distinct sources: first, a putative novel family of viruses; second, an archaeal virus (or plasmid); and third, mimivirus/mamavirus. The three genes shared with mimivirus/mamavirus were probably acquired by Sputnik after the association with APMV was established, and their products might be involved in the interaction of the virophage with its viral host. Within viral factories, recombination between the genomes of the virophage and APMV could result in an exchange of genes. APMV factories are probably capable of replicating foreign DNA, as suggested by experiments demonstrating efficient plasmid replication in poxvirus[10] and in African swine fever virus factories[11]. The presence of three genes homologous to mamavirus genes in the Sputnik genome suggests that gene transfer between Sputnik and mamavirus can occur during infection of *Acanthamoeba* by these two viruses together. It has been shown that some bacterial genes were recently acquired by mimivirus[12], but the source and the route of acquisition are still unknown[13]. Virophage could be a vehicle of such gene transfers, as well as of gene transfers between different giant viruses especially, if provirophages exist—a possibility that seems particularly plausible given the presence of genes for the predicted integrase and transposase subunit homologue in the virophage genome.

The integrase gene that is shared between Sputnik and archaeal viruses (plasmids) might have been independently derived from an ancestral virus that predated the divergence between archaea and

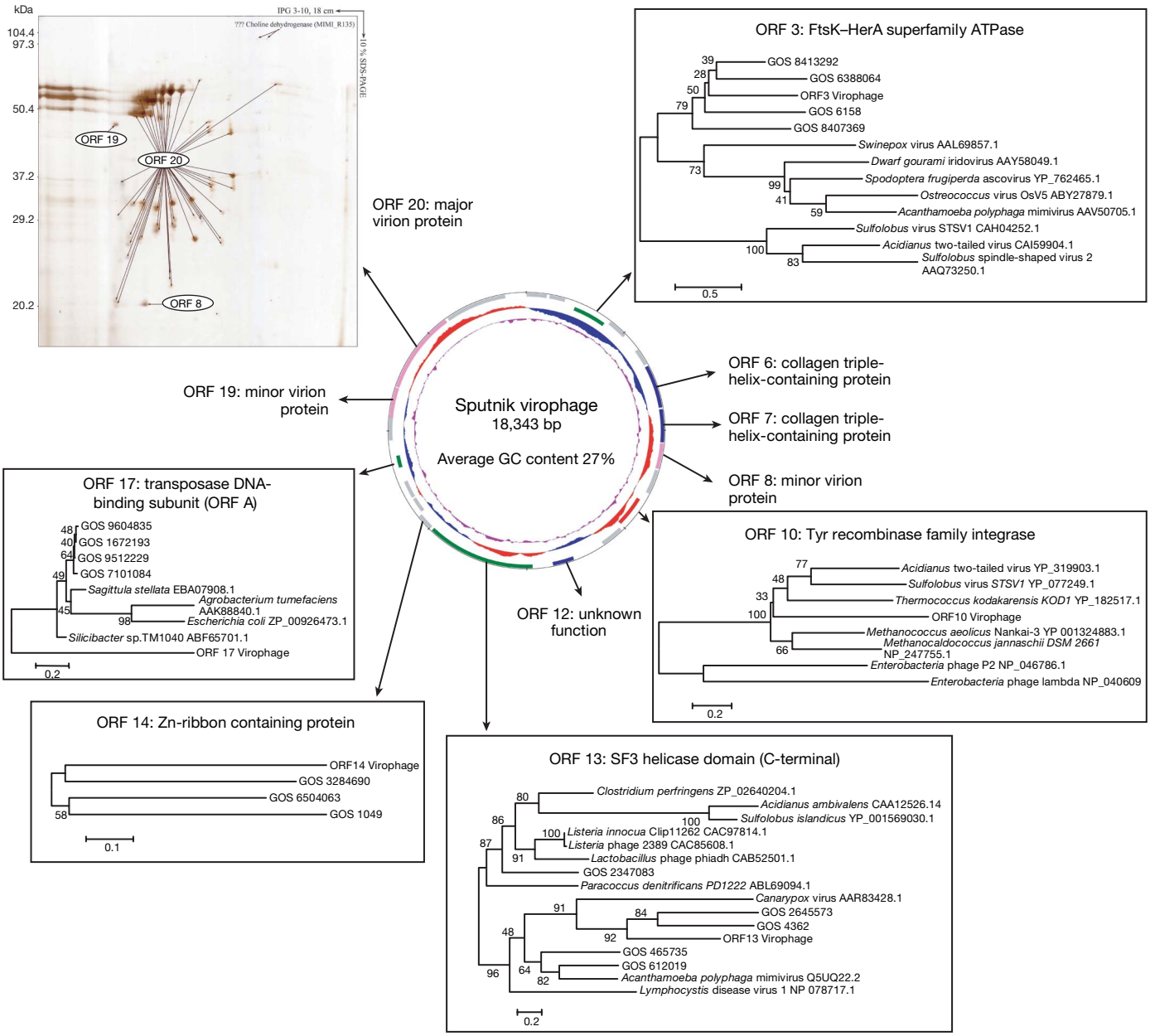

Figure 3 | The Sputnik chromosome. The predicted protein coding sequences are indicated on the two DNA strands (first, outer, circle) and coloured according to their corresponding homologues. ORFs with homologues to mamavirus/mimivirus are indicated in blue, ORFs with homologues to other NCLDVs and bacteriophages are shown in green, and the ORF homologous to an archaeal virus gene is shown in red. The virion protein coding sequences are shown in purple and ORFans are shown in grey. Phylogenetic trees are displayed for the predicted protein coding sequences with homologues in nr and/or the GOS data sets along with the 2D-gel identifying the capsid protein. GC skew and G + C content are indicated in the second and third circles, respectively. IPG, immobile pH-gradient buffer.

©2008 Macmillan Publishers Limited. All rights reserved

Table 1 | Homologies and predicted functions of the Sputnik protein coding sequences

Gene (size, amino-acid residues)	Closest homologue in GenBank nr (accession no., percentage identity/alignment length/*E*-value)	Closest homologue in the GOS data set (percentage identity/alignment length/*E*-value)	Domain architecture/protein family/predicted activity	Predicted function in virophage replication
ORF 1 (144)	–	–	Unknown	Unknown
ORF 2 (114)	–	–	Unknown	Unknown
ORF 3 (245)	RecA-superfamily ATPases (*Actinobacillus pleuropneumoniae* serovar 1 str. 4074) (ZP_00134596.2, 54%/35/0.01) MIMI L712	GOS_6857935 (48%/205/10^{-37})	FtsK–HerA superfamily ATPase	DNA packaging
ORF 4 (139)	Limited similarity to diverse Zn-ribbon proteins	–	Zn-ribbon-containing protein	Transcription regulation?
ORF 5 (119)	–	–	Unknown	Unknown
ORF 6 (310)	MIMI R196 (YP_142550.1, 53%/128/4×10^{-19})	GOS_3129237 (59%/130/10^{-23})	Collagen triple-helix-repeat-containing protein	Protein–protein interactions in factories
ORF 7 (236)	C1q and tumour necrosis factor related protein 5, mouse (NP_663588, 27%/156/0.001) MIMI R239	GOS_8448924 (57%/40/0.002)	Collagen triple-helix-repeat-containing protein	Protein–protein interactions in factories
ORF 8 (184)	–	–	Unknown	Minor virion protein
ORF 9 (175)	–	–	Unknown	Unknown
ORF 10 (226)	Phage integrase family protein (*Methanococcus aeolicus* Nankai-3) (YP_001324883, 32%/166/6×10^{-13})	–	Tyr recombinase family integrase	Integration of virophage into APMV genome?
ORF 11 (162)	–	–	Unknown	Unknown
ORF 12 (152)	MIMI R546 (Q5UR26, 64%/122/5×10^{-42})	–	Unknown	Unkown
ORF 13 (779)	Putative DNA-polymerase or DNA-primase (*Lactobacillus* phage phiadh) (NP_050131.1, 29%/171/4×10^{-12}) MIMI L207/206	Putative highly derived N-terminal primase domain, GOS_5022207 (32%/200/8×10^{-18}) C-terminal SF3 helicase domain GOS_2645573 (32%/409/4×10^{-46})	Primase–helicase	DNA replication
ORF 14 (114)	–	GOS_3284690 (45%/48/0.02)	Zn-ribbon-containing protein	Transcription regulation?
ORF 15 (109)	–	–	Membrane protein	Modification of APMV membrane?
ORF 16 (130)	–	–	Unknown	Unknown
ORF 17 (88)	–	GOS_9512229 (27%/80/0.03)	IS3 family transposase A protein	DNA-binding protein
ORF 18 (167)	–	–	Unknown	Unknown
ORF 19 (218)	–	–	Unknown	Minor virion protein
ORF 20 (595)	–	–	Unknown	Major capsid protein
ORF 21 (438)	–	-	Unknown	Unknown

eukaryotes. Alternatively, Sputnik might have acquired this gene from a virus (plasmid) harboured by an archaeal endosymbiont residing in a eukaryotic cell infected by Sputnik. Regardless of the exact source of this gene, one of the most remarkable features of the virophage is its apparent chimaeric origin. This seems to be one of the most convincing cases so far of gene mixing and matching within the virus world[14]. A search for additional virophages should shed more light on this unique mode of interaction between viruses.

As Sputnik multiplies in the APMV giant factories, it resembles satellite viruses of animals (for example adeno-associated viruses or hepatitis D virus) and plants (for example satellite tobacco necrosis virus)[15]. However, Sputnik reproduction seems to impair the production of normal APMV virions significantly, indicating that it is a genuine parasite. To our knowledge, this observation of a virus using the viral factory of another virus to propagate at the expense of its viral host has not been described previously. We have therefore termed this virus a virophage by analogy with bacteriophages; should other similar agents be discovered in the future, virophage could be used as a generic name to denote them.

METHODS SUMMARY

Isolation of viruses was performed on water sampled in a cooling tower as described previously[16]. For developmental cycle analysis, *A. castellanii* cells were infected with mamavirus alone or with Sputnik (Supplementary Information) and examined by transmission electronic microscopy and fluorescence as described previously for mimivirus[4].

Large volumes of *A. castellanii* infected by mamavirus and Sputnik were cultured. The culture supernatants were then filtered through 0.8-μm and 0.2-μm membranes. Sputnik particles were concentrated from the 0.2-μm filtrate, whereas mamavirus was obtained by washing the 0.2-μm membranes with K36 buffer. DNA was extracted by following the mimivirus procedure[1]. The genomes of the two viruses were sequenced on the 454-Roche GS20 as described[17]. Putative open reading frames were searched with GeneMark.hmm 2.0 (ref. 18), and translated sequences were compared with GenBank nr and the GOS data set (http://www.ncbi.nlm.nih.gov). MAFFT version 6 (ref. 19) or MUSCLE[20] was used to construct multiple alignments, and MEGA 4 (ref. 21) or TREEFINDER[22] was used to construct phylogenetic trees. Peptide data from excised spots were analysed by MALDI–TOF mass spectrometry as reported previously[23]. For western blot analysis, sera of BALB/c mice immunized with mamavirus or Sputnik were first absorbed on mimivirus and then on amoebae lysate.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 June; accepted 27 June 2008.
Published online 6 August 2008.

1. Raoult, D. *et al.* The 1.2-megabase genome sequence of Mimivirus. *Science* **306,** 1344–1350 (2004).
2. La Scola, B. *et al.* A giant virus in amoebae. *Science* **299,** 2033 (2003).
3. Koonin, E. V. Virology: Gulliver among the Lilliputians. *Curr. Biol.* **15,** R167–R169 (2005).
4. Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L. & Raoult, D. Ultrastructural characterization of the giant volcano-like virus factory of *Acanthamoeba polyphaga Mimivirus*. *PLoS ONE* **2,** e328 (2007).
5. Rusch, D. B. *et al.* The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **5,** e77 (2007).
6. Raoult, D. & Forterre, P. Redefining viruses: lessons from Mimivirus. *Nature Rev. Microbiol.* **6,** 315–319 (2008).
7. Rasmussen, M., Jacobsson, M. & Bjorck, L. Genome-based identification and analysis of collagen-related structural motifs in bacterial and viral proteins. *J. Biol. Chem.* **278,** 32313–32316 (2003).
8. Williamson, S. J. *et al.* The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. *PLoS ONE* **3,** e1456 (2008).
9. Iyer, L. M., Makarova, K. S., Koonin, E. V. & Aravind, L. Comparative genomics of the FtsK–HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res.* **32,** 5260–5279 (2004).

©2008 Macmillan Publishers Limited. All rights reserved

10. De Silva, F. S. & Moss, B. Origin-independent plasmid replication occurs in vaccinia virus cytoplasmic factories and requires all five known poxvirus replication factors. *Virol. J.* **2,** doi:10.1186/1743-422X-2-23 (2005).
11. Oliveira, S. & Costa, J. V. Replication of transfected plasmid DNA by cells infected with African swine fever virus. *Virology* **207,** 392–399 (1995).
12. Filee, J., Siguier, P. & Chandler, M. I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends Genet.* **23,** 10–15 (2007).
13. Moreira, D. & Brochier-Armanet, C. Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC Evol. Biol.* **8,** doi:10.1186/1471-2148-8-12 (2008).
14. Koonin, E. V., Senkevich, T. G. & Dolja, V. V. The ancient Virus World and evolution of cells. *Biol. Direct* **1,** doi:10.1186/1745-6150-1-29 (2006).
15. Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. (eds) *Virus Taxonomy* (*Eighth Report of the International Committee on Taxonomy of Viruses*) 1163–1169 (Elsevier, London, 2005).
16. La Scola, B., Barrassi, L. & Raoult, D. Isolation of new fastidious α-Proteobacteria and *Afipia felis* from hospital water supplies by direct plating and amoebal co-culture procedures. *FEMS Microbiol. Ecol.* **34,** 129–137 (2000).
17. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).
18. Lukashin, A. V. & Borodovsky, M. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* **26,** 1107–1115 (1998).
19. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30,** 3059–3066 (2002).
20. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).
21. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24,** 1596–1599 (2007).
22. Jobb, G., von Haeseler, A. & Strimmer, K. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4,** doi:10.1186/1471-2148-4-18 (2004).
23. Kowalczewska, M. & Raoult, D. Advances in *Tropheryma whipplei* research: the rush to find biomarkers for Whipple's disease. *Future Microbiol.* **2,** 631–642 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank X. de Lamballerie, S. Azza, P. de Clocquement, L. Espinosa, B. Campagna, N. Aldrovandi, V. Brice, A. Bernard, C. Ivars, B. Giumelli and Y. Wolf for expert assistance. This work was funded by the Centre National de la Recherche Scientifique (CNRS, crédits récurrents). I.P. is funded by a CIFFRE fellowship, E.K. is supported by the Intramural Research Program of the National Institutes of Health, National Library of Medicine, and P.F. is funded by the Institut Universitaire de France.

Author contributions D.R. and B.L.S. supervised the project and wrote the manuscript. C.D., P.F. and E.K. contributed to sequence analysis, interpretation of the results and writing of the manuscript. I.P. isolated the virus. M.S.-M. contributed to viral cycle analysis, interpretation of the results and writing of the manuscript. M.M. provided water samples. L.B. conducted the viral cycle experiment. C.R. and G.F. sequenced the genome.

Author Information The virophage genome has been deposited in GenBank under accession number EU606015. The *Acanthamoeba castellanii mamavirus* genes with homologues found in the Sputnik genome have been deposited in GenBank under accession numbers EU827539–EU827541. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.R. (didier.raoult@gmail.com).

©2008 Macmillan Publishers Limited. All rights reserved

METHODS

Inactivation of Sputnik. To obtain a pure suspension of mamavirus we proposed that, as observed previously for mimivirus[1], it would be resistant to high temperatures. We therefore subjected a supernatant containing Sputnik and mamavirus to 65 °C for 1 h. This suspension was then diluted in PAS (Page's amoebal saline) buffer by tenfold dilutions from 10^{-1} to 10^{-10}. Each dilution was inoculated into four culture wells of a suspension of fresh amoebae and observed daily for lysis under an inverted microscope. The last dilution producing lysis in one in four wells was 10^{-5}. The supernatant of this well was subcultured onto fresh amoebae, and an absence of Sputnik was verified by transmission electronic microscopy, immunofluorescence staining and Sputnik-specific PCR (see Supplementary Methods and Supplementary Results).

Evaluation of the effect of Sputnik on the developmental cycle of mamavirus. Supernatant containing Sputnik and mamavirus from infected *A. castellanii* was filtered through a 0.2-μm membrane and the Sputnik-containing filtrate was saved. A suspension of 10 ml of pure mamavirus was divided between two tubes. In tube 1, 200 μl of the Sputnik-containing supernatant was added. In tube 2, 200 μl of PAS buffer was added. *A. castellanii* cells (10 ml, 5×10^5 ml^{-1} in PAS buffer) were inoculated into culture flasks. In one flask, 1 ml of tube 1 was added; in a second flask, 1 ml of tube 2 was added, and 1 ml of PAS was added in the third flask. Living trophozoites were counted in each flask after 24 h. At 48 h after inoculation, mamavirus (flask 2) or Sputnik and mamavirus (flask 1) culture supernatants were used for titration of mamavirus and were then frozen. Titration was performed by endpoint dilution from 10^{-1} to 10^{-10} as described above and then with fivefold dilutions from 10^{-4} to 10^{-6}. Dilutions were scored until day 5 for lysis indicating mamavirus multiplication. The presence or absence of mamavirus multiplication was confirmed by detection with PCR in the supernatants from wells (data not shown).

To evaluate the effect of Sputnik on the appearance of abnormal mamavirus particles, monolayers of *A. castellanii* cells infected by mamavirus alone and by Sputnik and mamavirus were prepared for transmission electron microscopy. To normalize the comparison, counts of viral particles were performed in an area with a width of 1.5 μm around the virus factory limits.

Purification of viruses, preparation of viral DNA, and sequencing of Sputnik virus and mamavirus genomic DNA. Large volumes of *A. castellanii* cells infected by mamavirus and Sputnik were cultured. Viral supernatant were collected at 24–48 h, when lysis of amoebae was almost complete, by low-speed (100*g*) centrifugation for 15 min.

Sputnik was purified by filtration on 0.8-μm and 0.2-μm membranes. The filtrate was concentrated by ultracentrifugation at 100,000*g* for 70 min at 4 °C. The pellet was resuspended in K36 buffer, loaded on a 25% sucrose cushion in K36 and centrifuged with the same conditions. The purified pellet was washed once in K36 and resuspended in 10 mM Tris-HCl, 1 mM EDTA. To avoid contamination from DNA and RNA from amoebae, the suspension was treated twice with 10 μl of DNaseI_RNaseI-free (Roche) and 10 μl of RNaseI_DNaseI-free (Roche) and incubated for 60 min at 37 °C. The enzymes were inactivated by heating for 10 min at 70 °C. The DNA was extracted by following the mimivirus procedure[1]. A semiquantitative PCR was performed with primers specific for the 18S rRNA gene from amoebae[1] to estimate the contamination with DNA from amoebae. The Sputnik genome was pyrosequenced on 454–Roche GS20 as described[17]. The raw data were assembled by the gsAssembler of the GSFLX (35-bp overlap; 95% identity) leading to a large contig of 16.9 kilobases (kb) and four smaller contigs, for a total of 1.08 kb. Four primer sets were designed to close the molecule by PCR.

To obtain mamavirus DNA, the 0.2-μm membranes were washed with K36 buffer and this suspension was processed as above for sucrose density purification and for treatments with DNase/RNase. The pellet was then resuspended in TSD buffer (40 mM Tris-HCl pH 8, 2% SDS, 60 mM dithiothreitol) and incubated for 30 min at 60 °C with checking for lysis. If needed, an additional 25 μl of buffer was added to achieve total lysis, and this could be repeated three times. The suspension was diluted 1:10 in 50 mM Tris-HCl and treated with 10% Proteinase K at 56 °C. After three phenol/chloroform extractions, the DNA was precipitated with ethanol and resuspended in 75 μl of 10 mM Tris-HCl, 1 mM EDTA. The quality and the yield of the DNA was analysed on an agarose gel and stained with ethidium bromide. A semiquantitative PCR was performed with primers targeting the 18S rRNA gene from amoebae[1] to estimate contamination with DNA from amoebae. The mamavirus genome was also sequenced on 454–Roche GS20 and assembled with gsAssembler (40-bp overlap; 90% identity); 43 large contigs (more than 1.5 kb) were constructed for a genome size of 1.18 megabases. The average contig size was 27 kb; the largest was 173 kb. Taking into account all the contigs, 163 were obtained for a genome size of about 1.20 megabases.

Sequence analyses. Putative ORFs were defined with GeneMark.hmm 2.0 (ref. 18). Significant similarities of the ORF translated sequences were assessed through BLASTP and psi-BLAST[24] searches against the NCBI non-redundant protein database (http://www.ncbi.nlm.nih.gov). Functional motifs and conserved domains were identified by searches against PFAM version 22.0 (ref. 25), the Conserved Domain Database (CDD version 2.13), and SMART[26]. Homologues of Sputnik proteins in the environmental sequence data were detected by searching the NCBI environmental data set using BLASTP. Analyses of GC percentages and GC skew were performed with the online DNA Base Composition Analysis Tool (http://molbiol-tools.ca). The genome map was generated with Genomeviz[27]. MAFFT version 6 (ref. 19) or MUSCLE[20] was used to construct multiple alignments. Phylogenetic analyses were conducted with MEGA 4 (ref. 21) or TREEFINDER[22].

24. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
25. Finn, R. D. *et al.* Pfam: clans, web tools and services. *Nucleic Acids Res.* **34,** D247–D251 (2006).
26. Letunic, I. *et al.* SMART 4.0: towards genomic data integration. *Nucleic Acids Res.* **32,** D142–D144 (2004).
27. Ghai, R., Hain, T. & Chakraborty, T. GenomeViz: visualizing microbial genomes. *BMC Bioinformatics* **5,** doi:10.1186/1471-2105-5-198 (2004).

©2008 Macmillan Publishers Limited. All rights reserved