



Genome Sequences of New Faustovirus Strains ST1 and LC9, Isolated from the South of France

Amina Cherif Louazani, Julien Andreani, Maryem Ouarhache, Sarah Aherfi, Emeline Baptiste, Anthony Levasseur, Bernard La Scola

URMITE, Aix Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, IHU-Méditerranée Infection, Marseille, France

ABSTRACT Faustoviruses are amoeba-infecting giant viruses closely related to the *Asfarviridae* family. Here, we report the isolation, genome sequencing, and annotation of ST1 and LC9, two new strains belonging to lineages L and E9 of Faustoviruses, currently represented by only one representative each.

Faustoviruses are an expanding proposed family of nucleocytoplasmic large DNA viruses, or amoeba-infecting giant viruses, isolated on *Vermamoeba vermiformis* (1). Analyses of their genomes have revealed homology with *Asfarviridae* and Kaumoebavirus, suggesting a shared origin (2, 3). Here, we report the complete genome sequence of Faustovirus strain ST1 and the draft genome annotation of Faustovirus strain LC9. Faustovirus ST1 was isolated from wastewater in Saint-Pierre-de-Mézoargues, France, and strain LC9 was isolated from sewage in La Ciotat, France.

ST1 was isolated following the procedure previously described by Khalil et al. (4). It was produced and purified for genome sequencing on the Illumina MiSeq platform with a 2×250 -bp library in paired-end, parallel format. The genome was assembled with the Abyss program (5) and joined into a single scaffold with SSPACE (6). The LC9 draft genome sequences were retrieved from the European Nucleotide Archive database (WGS Sequence Set no. CZDJ01000000).

Subsequently, the analysis of both genomes followed the same procedure. tRNAs were predicted using tRNAscan-SE (7) and ARAGORN (8) software. Protein-coding genes were predicted using Genemarks (9). Predicted proteins above 50 amino acids in length and presenting at least one best hit in a BLASTp search against the NCBI nonredundant protein sequence database (E-value $> 1 \times 10^{-3}$), or above 99 amino acids in length and no BLASTp best hit, were considered for further analysis. For functional annotation, predicted proteins were searched for conserved domains and putative functions using the Pfam protein families database (10) and NCBI Conserved Domain Database search tools (11). For phylogenetic analysis, amino acid sequences were aligned using the MUSCLE program (12). The maximum likelihood statistical method with 1,000 bootstrap replicates using the Jones-Taylor-Thornton (JTT) model for amino acid substitution was used for tree construction on MEGA7 software (13).

Faustovirus ST1 has a total genome size of 470,659 bp and displays a G+C content of 36.7%. The five scaffolds comprising the Faustovirus LC9 genome have a total size of 470,227 bp and a mean G+C content of 39.8%. Results of tRNA prediction on both genomes showed the absence of tRNAs. The BLASTp and functional annotation resulted in 478 and 477 genes for ST1 and LC9, respectively. Among these genes, seven in ST1 and eight in LC9 present no homology on the NR database, constituting ORFans.

Phylogenetic analysis revealed that the new strains cluster with previously described lineages of Faustovirus (L for ST1 and E9 for LC9) (3). It is important to note that the ST1 and LC9 strains represent the second isolates in these lineages, which were previously

Received 16 May 2017 Accepted 18 May 2017 Published 13 July 2017

Citation Louazani AC, Andreani J, Ouarhache M, Aherfi S, Baptiste E, Levasseur A, La Scola B. 2017. Genome sequences of new Faustovirus strains ST1 and LC9, isolated from the south of France. *Genome Announc* 5:e00613-17. <https://doi.org/10.1128/genomeA.00613-17>.

Copyright © 2017 Louazani et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

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

represented by only one strain each (Faustovirus Liban for L and Faustovirus E9 for E9). These results suggest that similar strains are spread in separate environments worldwide: Faustovirus Liban being isolated from seawater in Lebanon and ST1 from house wastewater in France. This observation is in concordance with previously reported results for *Marseilleviridae* and *Mimiviridae* spp. (14–16). Moreover, Faustoviruses are recurrently isolated from sewage and wastewater, suggesting the presence of a potential reservoir in close human proximity.

Accession number(s). The complete annotated genome sequences of Faustovirus strains ST1 and LC9 have been deposited in ENA under the accession numbers [LT839607](https://doi.org/10.1093/nar/gkv1344) and [CZDJ02000000](https://doi.org/10.1093/nar/gkh454), respectively.

REFERENCES

1. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, Rossmann M, Colson P, Raoult D, La Scola B. 2015. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J Virol* 89:6585–6594. <https://doi.org/10.1128/JVI.00115-15>.
2. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, Raoult D, La Scola B. 2016. Kaumobavirus, a new virus that clusters with faustoviruses and *Asfarviridae*. *Viruses* 8:278. <https://doi.org/10.3390/v8110278>.
3. Benamar S, Reteno DGI, Bandy V, Labas N, Raoult D, La Scola B. 2016. Faustoviruses: comparative genomics of new *Megavirales* family members. *Front Microbiol* 7:3. <https://doi.org/10.3389/fmicb.2016.00003>.
4. Bou Khalil JY, Andreani J, Raoult D, La Scola B. 2016. A rapid strategy for the isolation of new faustoviruses from environmental samples using *Vermamoeba vermiformis*. *J Vis Exp*. <https://doi.org/10.3791/54104>.
5. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res* 19:1117–1123. <https://doi.org/10.1101/gr.089532.108>.
6. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
7. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:W686–W689. <https://doi.org/10.1093/nar/gki366>.
8. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <https://doi.org/10.1093/nar/gkh152>.
9. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618. <https://doi.org/10.1093/nar/29.12.2607>.
10. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. <https://doi.org/10.1093/nar/gkv1344>.
11. Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 32:W327–W331. <https://doi.org/10.1093/nar/gkh454>.
12. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
13. Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
14. Pagnier I, Reteno DG, Saadi H, Boughalmi M, Gaia M, Slimani M, Ngounga T, Bekliz M, Colson P, Raoult D, La Scola B. 2013. A decade of improvements in *Mimiviridae* and *Marseilleviridae* isolation from amoeba. *Intervirology* 56:354–363. <https://doi.org/10.1159/000354556>.
15. Doutre G, Arfib B, Rochette P, Claverie JM, Bonin P, Abergel C. 2015. Complete genome sequence of a new member of the *Marseilleviridae* recovered from the brackish submarine spring in the cassis Port-Miou Calanque, France. *Genome Announc* 3(6):e01148–15. <https://doi.org/10.1128/genomeA.01148-15>.
16. Yoosuf N, Yutin N, Colson P, Shabalina SA, Pagnier I, Robert C, Azza S, Klose T, Wong J, Rossmann MG, La Scola B, Raoult D, Koonin EV. 2012. Related giant viruses in distant locations and different habitats: *Acanthamoeba polyphaga moumouvirus* represents a third lineage of the *Mimiviridae* that is close to the *Megavirus* lineage. *Genome Biol Evol* 4:1324–1330. <https://doi.org/10.1093/gbe/evs109>.