Genome Sequence of a Novel Archaeal Fusellovirus Assembled from the Metagenome of a Mexican Hot Spring
Servin-Garciduenas, L. E.; Peng, Xu; Garrett, Roger Antony; Martinez-Romero, E.

Published in:
Genome Announcements

DOI:
10.1128/genomeA.00164-13

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Genome Sequence of a Novel Archaeal Fusellovirus Assembled from the Metagenome of a Mexican Hot Spring

Luis E. Servín-Garcidueñas, Xu Peng, Roger A. Garrett, Esperanza Martínez-Romero
Department of Ecological Genomics, Center for Genomic Sciences, National University of Mexico, Cuernavaca, Morelos, Mexico; Department of Biology, Archaea Centre, University of Copenhagen, Copenhagen, Denmark

The consensus genome sequence of a new member of the family Fuselloviridae designated as SMF1 (Sulfolobales Mexican fusellovirus 1) is presented. The complete circular genome was recovered from a metagenomic study of a Mexican hot spring. SMF1 exhibits an exceptional coding strand bias and a reduced set of fuselloviral core genes.

Members of the Fuselloviridae family from the crenarchaeal order Sulfolobales have been characterized, and they are abundant in extreme geothermal environments (1, 2). They carry circular double-stranded DNA (dsDNA) genomes and exhibit spindle-shaped morphologies. Here, we report the consensus genome sequence of a novel fusellovirus recovered from aqueous sediments from Los Azufres, Mexico.

Samples were collected from a hot spring with a pH of 3.6 and a temperature of 65°C. DNA was purified using the UltraClean microbial and the UltraClean Mega soil DNA kits (MoBio Laboratories, Inc., Carlsbad, CA). Sequencing was performed on an Illumina GAIIx platform, producing 36-bp paired-end reads with 300-bp inserts representing 216 Mb. Reads were assembled using Velvet 1.2.07 (3). A set of contigs were predicted by BLASTX searches to be of fuselloviral origin. Gaps were closed iteratively by mapping and reassembling reads to these contigs using Maq 0.7.1 (4) and Velvet. Open reading frames (ORFs) were predicted using GeneMark.hmm2.0 (5) and were manually verified using Artemis (6).

The average sequence coverage of the 14,847-bp circular dsDNA genome was 1,257-fold. We detected 57 candidate single nucleotide polymorphisms by Maq. The G+C content was 45.43%, higher than the 37.5 to 39.7% content of other fuselloviruses. The SMF1 genome shows exceptional coding-strand bias and a reduced set of fuselloviral core genes.

In conclusion, the SMF1 genome was recovered from a site widely separated geographically from the locations of other sequenced fuselloviruses. The SMF1 genome shows exceptional properties, including a coding-strand bias and a high incidence of genes organized in operon structures, but nevertheless, it retains a large set of conserved fusellovirus genes, which lends further support to the exchange of genetic material over intercontinental distances (2, 10).

Nucleotide sequence accession number. The genome sequence was deposited in GenBank under the accession no. KC618393.

ACKNOWLEDGMENTS

This research was supported by PAPIIT IN205412 from DGAPA, UNAM, and by SUBNARGEM, SAGARPA. L.E.S.G. received a PhD scholarship from CONACYT (Mexico). L.E.S.-G. had a “Beca Mixta” from CONACYT and a fellowship from PAEP during a research internship at the Danish Archaea Center.

We thank the Programa de Doctorado en Ciencias Biomédicas from UNAM, the UUSM from UNAM for sample sequencing, and Jesus Campos Garcia from the UMSNH for providing laboratory facilities. The CFE personnel provided a permit for samplings.
REFERENCES