

Genome Sequence of a Novel Archaeal Rudivirus Recovered from a Mexican Hot Spring

Luis E. Servín-Garcidueñas,^a Xu Peng,^b Roger A. Garrett,^b Esperanza Martínez-Romero^a

Department of Ecological Genomics, Center for Genomic Sciences, National University of Mexico, Cuernavaca, Morelos, Mexico^a; Department of Biology, Archaea Centre, University of Copenhagen, Copenhagen, Denmark^b

We report the consensus genome sequence of a novel GC-rich rudivirus, designated SMR1 (*Sulfolobales* Mexican rudivirus 1), assembled from a high-throughput sequenced environmental sample from a hot spring in Los Azufres National Park in western Mexico.

Received 12 October 2012 Accepted 29 October 2012 Published 15 January 2013

Citation Servín-Garcidueñas LE, Peng X, Garrett RA, Martínez-Romero E. 2013. Genome sequence of a novel archaeal rudivirus recovered from a Mexican hot spring. *Genome Announc.* 1(1):e00040-12. doi:10.1128/genomeA.00040-12.

Copyright © 2013 Servín-Garcidueñas et al. This is an open-access article distributed under the terms of the [Attribution 3.0 Unported Creative Commons License](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Luis E. Servín-Garcidueñas, luis.e.servin@gmail.com.

Thermophilic archaeal viruses have been isolated from thermal terrestrial sites, revealing an incredibly large viral diversity (1–3). Mexico contains active volcanoes and geothermal areas extending along the Trans-Mexican Volcanic Belt (TMVB) (4). However, the thermophilic viral diversity present within the TMVB hot springs remains unexplored. Here, we present the consensus genome sequence of a novel rudivirus recovered by iterative *de novo* read mapping and assembly from the metagenome of a hot spring located along the northern edge of the TMVB.

Aqueous sediment samples were collected from an acidic hot spring (pH 3.6 and 65°C) located at Los Azufres, Mexico, in March 2009. The DNA was purified using the UltraClean microbial DNA and UltraClean Mega soil DNA kits (MoBio Laboratories, Inc., Carlsbad and Solana Beach, CA). The metagenomic DNA was sequenced with an Illumina GAIx platform producing 36-bp paired-end reads with 300-bp inserts representing 216 Mbp. The reads were assembled *de novo* using Velvet 1.2.07 (5). The contigs with overrepresented coverage were verified by BLASTX searches to be of viral origins. The reads were mapped to the viral contigs using Maq 0.7.1 (6), and the mapping reads were reassembled to eliminate gaps. The coding sequences were predicted using GeneMark.hmm 2.0 (7) and were manually verified using Artemis (8).

The sequence coverage of the 27,431-bp double-stranded DNA genome was 240-fold. The presence of an inverted terminal repeat (1,240 bp), characteristic of the linear rudiviral genomic termini, indicated that the genome was complete or almost complete. The G+C content of 46.6% was higher than the 25 to 39% content of the four rudiviral genomes characterized previously (9–11). The host is likely to be a member of the order *Sulfolobales*, the sequences of which dominated the metagenome. Moreover, the G+C content of *Sulfolobales* Mexican rudivirus 1 (SMR1) was similar to those of the *Metallosphaera* genomes (~45%). Thirty-seven open reading frames (ORFs) were identified, 19 of which have putative homologs in the other characterized rudiviruses; this strongly supports SMR1 being a member of the *Rudiviridae* family. Common annotated gene products include the major coat protein, three minor structural proteins, two glycosyl transferases,

a clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas4-like protein, a putative replication protein, a Holliday junction helicase, a Holliday junction resolvase, an S-adenosylmethionine-dependent methyltransferase, and a putative transcriptional regulator. The seven other shared rudiviral proteins were not assigned functions.

Three additional ORFs show sequence similarities to archaeal ORFs, including one carrying a zinc finger SWIM domain and a predicted CopG domain. Four other ORFs contain domains related to the thioredoxin-like superfamily and the GTP-binding proteins, as well as a ribbon-helix-helix protein and a nop25 domain-containing protein. Eight additional ORFs showed no significant matches. Interestingly, three ORFs were related to viral ORFs of the *Lipothrixviridae* family. Of the 6,000,792 environmental reads, 183,365 (3.05%) mapped to the consensus viral genome and 115 candidate single nucleotide polymorphisms (SNPs) were detected by Maq.

In conclusion, despite the large geographical distance from the locations of other sequenced rudiviruses, SMR1 retained a core set of conserved rudiviral genes that were inferred to be important for the viral life cycle.

Nucleotide sequence accession number. The genome sequence was deposited in GenBank under the accession no. [JX944686](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/seqview.fcgi?acc=JX944686).

ACKNOWLEDGMENTS

The UUSM from UNAM is thanked for sample sequencing. The Comisión Federal de Electricidad personnel provided a permit for samplings. We thank professor Jesús Campos García from the UMSNH for providing laboratory facilities. The samplings were carried out with the efforts of Hans Román, Jonathan Lopez, José Luis Servín, and Cecilia Garcidueñas.

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

REFERENCES

1. Lawrence CM, Menon S, Eilers BJ, Bothner B, Khayat R, Douglas T, Young MJ. 2009. Structural and functional studies of archaeal viruses. *J. Biol. Chem.* **284**:12599–12603.
2. Pina M, Bize A, Forterre P, Prangishvili D. 2011. The archeoviruses. *FEMS Microbiol. Rev.* **35**:1035–1054.
3. Prangishvili D, Forterre P, Garrett RA. 2006. Viruses of the Archaea: a unifying view. *Nat. Rev. Microbiol.* **4**:837–848.
4. Ferrari L, Orozco-Esquivel MT, Manea V, Manea M. 2011. The dynamic history of the trans-Mexican volcanic belt and the Mexico subduction zone. *Tectonophysics* **522–523**:122–149.
5. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**:821–829.
6. Li H, Ruan J, Durbin R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**: 1851–1858.
7. Besemer J, Borodovsky M. 1999. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res.* **27**:3911–3920.
8. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison Tool. *Bioinformatics* **21**: 3422–3423.
9. Peng X, Blum H, She Q, Mallok S, Brügger K, Garrett RA, Zillig W, Prangishvili D. 2001. Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. *Virology* **291**:226–234.
10. Vestergaard G, Häring M, Peng X, Rachel R, Garrett RA, Prangishvili D. 2005. A novel rudivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* **336**:83–92.
11. Vestergaard G, Shah SA, Bize A, Reitberger W, Reuter M, Phan H, Briegel A, Rachel R, Garrett RA, Prangishvili D. 2008. *Stygiolobus* rod-shaped virus and the interplay of crenarchaeal rudiviruses with the CRISPR antiviral system. *J. Bacteriol.* **190**:6837–6845.