UCSF
UC San Francisco Previously Published Works

Title

Draft Genome Sequence of Tombunodavirus UC1

Permalink

https://escholarship.org/uc/item/3pf562hv

Journal

Microbiology Resource Announcements, 3(4)

ISSN

2576-098X

Authors

Greninger, Alexander L

DeRisi, Joseph L

Publication Date

2015-08-27

DOI

10.1128/genomea.00655-15

Draft Genome Sequence of Tombunodavirus UC1

Alexander L. Greninger,[a,b] Joseph L. DeRisi[a,b]

Department of Biochemistry and Biophysics, UCSF, San Francisco, California, USA[a]; Howard Hughes Medical Institute, UCSF, San Francisco, California, USA[b]

We report here the draft genome sequence of tombunodavirus UC1 assembled from metagenomic sequencing of organisms in San Francisco wastewater. This virus shares hallmarks of members of the *Tombusviridae* and the nodavirus-like *Plasmopara halstedii* and *Sclerophthora macrospora* viruses.

Address correspondence to Alexander L. Greninger, gerbix@gmail.com.

*T*ombusviridae and *Nodaviridae* are both families of single-stranded RNA viruses that demonstrate a wide range of host organisms (1, 2). While performing weekly metagenomic sequencing of organisms in San Francisco wastewater, we assembled a contig of 4,244 nucleotides that by BLASTx aligned 35% at the amino acid level to the RNA-dependent RNA polymerase of the tombusviruses olive latent virus 1 and tobacco necrosis virus and 57% at the amino acid level to the coat protein of the nodavirus-like *Plasmopara halstedii* virus A (3). In addition to these alignments, the first portion of the contig aligned to pfam08500, representing the p33 replication accessory protein from the *Tombusviridae* family (4). The genome size of 4,244 nucleotides is consistent with that of other members of the *Tombusviridae* family, which range from 3.6 kb to 4.8 kb.

We have given the name tombunodavirus UC1 to this virus due to its sequence identity to members of both the nodavirus and tombusvirus families, although it appears the viral genome is monopartite, like a tombusvirus. The contig contains three separate open reading frames (ORFs) in the standard genetic code of 888, 1,482, and 1,353 nucleotides with an overlap of 29 nucleotides, while in a more permissive ciliate genetic code, as the dominant host organism present in the sample, it contains two ORFs of 2,493 and 1,626 nucleotides, representing the unique origins of the two ORFs, which overlap by 164 nucleotides. Members of the *Tombusviridae* family infect plants, while members of the nodavirus family infect a broad array of organisms, including fish and insects. The host organism of tombunodavirus UC1 is currently unknown.

This viral genome was recovered from a wastewater sample from 25 January 2010 that was taken after a large rainstorm left >5 in. of rain over the preceding week. This same sample also contained novel ciliate and marine RNA viruses and phages (5–7). Sample processing was performed on 1 liter of wastewater that was concentrated to <5 ml, with particles between the size of 0.22 μm and 300 kDa using Millipore Pellicon XL 300-kDa filters and 0.22-μm spin columns. Nucleic acid was extracted using the Zymo Viral DNA/RNA kit, and half of the recovered nucleic acid was treated with DNase. The contig was discovered and assembled using PRICE version 1.0 and SURPI version 1.0 from a total of 15,719,690 paired-end 65-bp reads sequenced on an Illumina GAIIx split between these DNAsed and untreated nucleic acid preparations (8, 9). The average coverage of the tombunodavirus contig was 5,089× with coverage 2.4× higher in the DNAsed library than that in the untreated library, consistent with an RNA genome.

Nucleotide sequence accession number. The GenBank accession number for tombunodavirus UC1 is KF510030.

REFERENCES

1. **White KA, Nagy PD**. 2004. Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. Prog Nucleic Acid Res Mol Biol **78**:187–226. http://dx.doi.org/10.1016/S0079-6603(04)78005-8.
2. **Schneemann A, Reddy V, Johnson JE**. 1998. The structure and function of nodavirus particles: a paradigm for understanding chemical biology, p 381–446. *In* Maramorosch K, Murphy F, Shatkin A (ed), Advances in virus research. Academic Press, San Diego, CA.
3. **Heller-Dohmen M, Göpfert JC, Pfannstiel J, Spring O**. 2011. The nucleotide sequence and genome organization of *Plasmopara halstedii* virus. Virol J **8**:123. http://dx.doi.org/10.1186/1743-422X-8-123.
4. **Nagy PD, Barajas D, Pogany J**. 2012. Host factors with regulatory roles in tombusvirus replication. Curr Opin Virol **2**:691–698. http://dx.doi.org/10.1016/j.coviro.2012.10.004.
5. **Greninger AL, DeRisi JL, Knowles S, Ng VL, Chaturvedi V**. 2015. Draft genome sequences of ciliovirus and brinovirus from San Francisco wastewater. Genome Announc **3**(3):e00651-15. http://dx.doi.org/10.1128/genomeA.00651-15.
6. **Greninger AL, DeRisi JL, Knowles S, Ng VL, Chaturvedi V**. 2015. Draft genome sequences of marine RNA viruses SF-1, SF-2, and SF-3 recovered from San Francisco wastewater. Genome Announc **3**(3):e00653-15. http://dx.doi.org/10.1128/genomeA.00653-15.
7. **Greninger AL, DeRisi JL, Knowles S, Ng VL, Chaturvedi V**. 2015. Draft genome sequences of *Leviviridae* RNA phages EC and MB recovered from San Francisco wastewater. Genome Announc **3**(3):e00652-15. http://dx.doi.org/10.1128/genomeA.00652-15.
8. **Ruby JG, Bellare P, Derisi JL**. 2013. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. G3 (Bethesda) **3**:865–880. http://dx.doi.org/10.1534/g3.113.005967.
9. **Naccache SN, Federman S, Veeraraghavan N, Zaharia M, Lee D, Samayoa E, Bouquet J, Greninger AL, Luk K-C, Enge B, Wadford DA, Messenger SL, Genrich GL, Pellegrino K, Grard G, Leroy E, Schneider BS, Fair JN, Martínez MA, Isa P, Crump JA, DeRisi JL, Sittler T, Hackett J, Miller S, Chiu CY**. 2014. A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. Genome Res **24**:1180–1192. http://dx.doi.org/10.1101/gr.171934.113.