

Draft Genome Sequences of *Leviviridae* RNA Phages EC and MB Recovered from San Francisco Wastewater

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 sequences of marine RNA phages EC and MB assembled from metagenomic sequencing of organisms in San Francisco wastewater. These phages showed moderate translated amino acid identity to other enterobacteria phages and appear to constitute novel members of the *Leviviridae* family.

Received 19 May 2015 Accepted 22 May 2015 Published 25 June 2015

Citation Greninger AL, DeRisi JL. 2015. Draft genome sequences of *Leviviridae* RNA phages EC and MB recovered from San Francisco wastewater. *Genome Announc* 3(3): e00652-15. doi:10.1128/genomeA.00652-15.

Copyright © 2015 Greninger and DeRisi. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

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

Leviviridae is a family of positive-stranded RNA bacteriophages consisting of two genera, *Allolevivirus* and *Levivirus*, which are each generally restricted to a single host genus of Gram-negative bacteria (1). Of note, the RNA-binding coat protein of the *Levivirus* MS2 phage has served as a tool for many basic science applications and as a model for the assembly of RNA viruses (2, 3). Interestingly, surveys performed in 1976 and 2006 of marine/coastal RNA virus communities found very few if any RNA phages (4).

While performing weekly metagenomic sequencing of organisms in San Francisco wastewater, we assembled 3,180-nucleotide (EC) and 3,925-nucleotide (MB) contigs at 849× and 229× coverage, respectively, which aligned by BLASTx to enterobacterial phages Hgal1, M11, and Qbeta. Each contig comprised a typical *Leviviridae* genome organization, consisting of three open reading frames (ORFs) encoding maturation, coat, and replicase proteins, with phage EC likely truncated by ~600 nucleotides at the 3' end. However, neither phage encoded a readily apparent lysis or read-through protein. Phage EC appears to possess ORFs that encode both coat and replicase proteins but lacks any protein with alignment to known lysis proteins of *Levivirus*. Phage MB also appears to lack a recognizable lysis protein.

The predicted replicase protein of phage MB demonstrated 39% amino acid identity to *Caulobacter* phage phiCb5 over the entirety of the replicase protein, while the closest replicase to phage EC demonstrated 42% amino acid identity to enterobacterial phage Hgal1. Both replicases demonstrated <38% identity to members of *Leviviridae* already placed in a genus, which is equivalent to amino acid identity between *Allolevivirus* and *Levivirus* genera, consistent with these two viruses forming two novel genera within *Leviviridae*. The predicted maturation proteins aligned with 32% amino acid identity to enterobacterial phage BZ13 (EC) and 28% to *Caulobacter* phage phiCb5 (MB). The predicted phage EC coat protein aligned with 32% amino acid identity to enterobacterial phage C-1 INW-2012, while the predicted phage MB coat protein failed to demonstrate significant amino acid alignments to any phage proteins but did significantly align by HHPred to *Caulobacter* phiCb5 virus-like particle (5, 6).

The sample that yielded these draft genome sequences was prepared from 1 liter of wastewater that was concentrated to <5 ml of 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 phage genomes were discovered and assembled using PRICE version 1.0, Geneious version 8.0 Assembler, 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 (7, 8).

Nucleotide sequence accession numbers. The GenBank accession numbers for marine RNA phages EC and MB are [KF616862](https://www.ncbi.nlm.nih.gov/nuccore/KF616862) and [KF510034](https://www.ncbi.nlm.nih.gov/nuccore/KF510034), respectively.

ACKNOWLEDGMENTS

This work was supported by the Howard Hughes Medical Institute and a gift from the Sergei Brin Foundation.

We thank Michael Kellogg at the Oceanside Wastewater Treatment Plant for his assistance.

REFERENCES

- Bollback JP, Huelsenbeck JP. 2001. Phylogeny, genome evolution, and host specificity of single-stranded RNA bacteriophage (family *Leviviridae*). *J Mol Evol* 52:117–128.
- Keryer-Bibens C, Barreau C, Osborne HB. 2008. Tethering of proteins to RNAs by bacteriophage proteins. *Biol Cell* 100:125–138. <http://dx.doi.org/10.1042/BC20070067>.
- Dykeman EC, Grayson NE, Toropova K, Ranson NA, Stockley PG, Twarock R. 2011. Simple rules for efficient assembly predict the layout of a packaged viral RNA. *J Mol Biol* 408:399–407. <http://dx.doi.org/10.1016/j.jmb.2011.02.039>.
- Hidaka T, ヒダカトミオ, 日高富男. 1975. Identification of the type of nucleic acid in marine bacteriophages with acridine orange staining. *Memoirs Fac Fish Kagoshima Univ* 24:133–138.
- Hildebrand A, Remmert M, Biegert A, Söding J. 2009. Fast and accurate automatic structure prediction with HHpred. *Proteins* 77(Suppl 9): 128–132. <http://dx.doi.org/10.1002/prot.22499>.
- Plevka P, Kazaks A, Voronkova T, Kotelovica S, Dishlers A, Liljas L, Tars K. 2009. The structure of bacteriophage phiCb5 reveals a role of the RNA

- genome and metal ions in particle stability and assembly. *J Mol Biol* 391: 635–647. <http://dx.doi.org/10.1016/j.jmb.2009.06.047>.
7. 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>.
 8. 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>.