genomeA
Journals.ASM.org

Draft Genome Sequences of Ciliovirus and Brinovirus 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 the draft genome sequences of ciliovirus and brinovirus, two members of a likely new family of RNA viruses assembled from San Francisco wastewater. Based on sequence alignments and a nonuniversal genetic code, we believe these to be the first described RNA viruses of ciliates; however, more work is necessary to confirm their host.

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

Citation Greninger AL, DeRisi JL. 2015. Draft genome sequences of ciliovirus and brinovirus from San Francisco wastewater. Genome Announc 3(3):e00651-15. doi:10.1128/genomeA.00651-15.

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

Ciliates are large single-cellular protozoal organisms defined by the presence of cilia as well as macro- and micronucleus. Ciliates are also noted for possessing a nonuniversal genetic code. Virus-like particles have been reported to be associated with *Hyalophysa chattoni*, and a large DNA virus, chlorella virus, has been isolated from *Paramecium bursaria* (1, 2). However, no RNA virus has been recovered from ciliates to date (3).

While performing weekly metagenomic sequencing of San Francisco wastewater, we recovered two contigs of 10,381 and 9,565 nucleotides that distantly aligned by BLASTx to a small portion of an RNA-dependent RNA polymerase of various members of the *Flaviviridae*. Translation of the two contigs using the universal genetic code failed to reveal open-reading frames (ORFs) of >500 nucleotides. However, translation of two contigs utilizing the ciliate genetic code revealed ORFs of 9,954 and >9,260 nucleotides that preserved the BLASTx alignments. HHPred analysis of the contigs translated in the ciliate genetic code revealed a superfamily II helicase similar to eIF4A followed by a picornavirus/flavivirus-like RNA-dependent RNA polymerase followed by a birnavirus/sobemovirus-like viral capsid protein, suggesting a genome organization most similar to that of the *Potyviridae* (3, 4). Of note, the two polyproteins were 28.7% identical by amino acid to each other throughout the ORF and 40% identical by amino acid to each other in the putative RNA-dependent RNA polymerase and helicase regions, consistent with these two contigs forming two new genera of a new viral family.

Notably, both contigs were recovered in only one wastewater sample from 25 January 2010 that was taken after a large rainstorm that left >5 inches of rain over the preceding week. Over one-quarter of the alignable nonchordate eukaryotic reads from the DNAsed sample aligned to *Ciliophora* organisms with half of these reads belonging to the *Tetrahymenidae* family, making it one of the most abundant organisms in the sample and strengthening the case for these viruses having ciliate hosts (5).

Sample processing was performed on 1 liter of wastewater that was concentrated to <5 mL with particles between the sizes of 0.22 μm and 300 kDa using Millipore Pellicon XL 300-kDa filters and 0.22-μm spin columns. Viral particle-enriched sample was treated with micrococcal nuclease and nucleic acid was extracted using a Zymo Viral DNA/RNA kit and half of the recovered nucleic acid was treated with DNase. Both contigs were assembled using PRICE v1.0 and Geneious v8.0 Assembler with metagenomic analysis using SURPI v1.0 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 (5, 6). Average coverage of the two contigs was 550× and 1271×, respectively. Coverage of ciliovirus in the DNAsed sample was 3.5× higher than in the untreated sample, while coverage of brinovirus was 2.4× higher in the DNase-treated sample, consistent with these being RNA viruses.

Nucleotide sequence accession numbers. The GenBank accession numbers for ciliovirus and brinovirus are JN661159 and KF412899, 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

1. **Kucera F.** 1992. Virus-like particles associated with the apostome ciliate *Hyalophysa chattoni*. Dis Aquat Org **12:**151–153. http://dx.doi.org/10.3354/dao012151.
2. **Van Etten JL, Meints RH, Kuczmarski D, Burbank DE, Lee K.** 1982. Viruses of symbiotic chlorella-like algae isolated from *Paramecium bursaria* and *Hydra viridis*. Proc Natl Acad Sci USA **79:**3867–3871. http://dx.doi.org/10.1073/pnas.79.12.3867.
3. **Koonin EV, Wolf YI, Nagasaki K, Dolja VV.** 2008. The big bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat Rev Microbiol **6:**925–939. http://dx.doi.org/10.1038/ nrmicro2030.
4. **Hildebrand A, Remmert M, Biegert A, Söding J.** 2009. Fast and accurate

automatic structure prediction with HHpred. Proteins **77:**128–132. http://dx.doi.org/10.1002/prot.22499.

5. **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.
6. **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.

genomeA
Journals.ASM.org

Draft Genome Sequences of Marine RNA Viruses SF-1, SF-2, and SF-3 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 the draft genome sequences of marine RNA viruses SF-1, SF-2, and SF-3, which were assembled from metagenomic sequencing of organisms in San Francisco wastewater. These viruses were most closely related to marine RNA virus JP-B and algae viruses.

Received 19 May 2015 **Accepted** 20 May 2015 **Published** 18 June 2015

Citation Greninger AL, DeRisi JL. 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. doi:10.1128/genomeA.00653-15.

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

The picorna-like superfamily is a rapidly expanding taxonomic unit of positive-stranded RNA viruses with conserved RNA-dependent RNA polymerase (RdRp), capsid, and helicase proteins that have a broad host range, including animals, plants, and insects (1–3). While performing weekly metagenomic sequencing of organisms in San Francisco wastewater, we assembled three contigs of 8,695, 9,270, and 8,642 nucleotides that aligned by BLASTx to the RNA-dependent RNA polymerase and capsid genes of the marine RNA virus JP-B (30 to 40% amino acid identity), a member of the *Picornavirales* order (1, 2). The first two contigs contained bicistronic viral genomes with open reading frames (ORFs) of 4,782/2,877 nucleotides (RNA-dependent RNA polymerase-containing polyprotein ORF/structural polyprotein ORF) and 5,178/2,664 nucleotides, while the third genome contained a single ORF of 7,731 nucleotides in the standard genetic code. The ORFs aligned by 30 to 38% to each other and ~30% by amino acids to RNA viruses of *Chaetoceros tenuissimus*, *Rhizosolenia setigera*, and *Asterionellopsis glacialis*, suggesting that these may be RNA viruses of algae.

All three contigs were assembled from a single metagenomic library derived from a wastewater sample taken on 25 January 2010 following a large rainstorm that left >5 inches of rain over the preceding week. Unlike the likely ciliate viruses also discovered in this sample, wastewater samples collected in March 2010 also contained reads to these viruses but at a significantly lower sequence count. This sample was created by 200-fold concentration of 1 liter of wastewater, 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. The viral particle-enriched sample was treated with micrococcal nuclease, nucleic acid was extracted using the Zymo viral DNA/RNA kit, and half of the recovered nucleic acid was treated with DNase. The three contigs were discovered and assembled using PRICE version 1.0 (4), the Geneious version 8.0 Assembler, and SURPI version 1.0 (5) 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 (3, 6). The average coverages of the three contigs using all reads from the sample were 1,749×, 1,379×, and 222×.

Nucleotide sequence accession numbers. The GenBank accession numbers for marine RNA viruses SF-1, SF-2, and SF-3 are JN661160, KF412901, and KF478836, 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

1. **Koonin EV, Wolf YI, Nagasaki K, Dolja VV**. 2008. The big bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat Rev Microbiol **6**:925–939. http://dx.doi.org/10.1038/nrmicro2030.
2. **Greninger AL**. 2015. *Picornavirus*-host interactions to construct viral secretory membranes. Prog Mol Biol Transl Sci **129**:189–212. http://dx.doi.org/10.1016/bs.pmbts.2014.10.007.
3. **Culley AI, Lang AS, Suttle CA**. 2003. High diversity of unknown picorna-like viruses in the sea. Nature **424**:1054–1057. http://dx.doi.org/10.1038/nature01886.
4. **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.
5. **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.
6. **Culley AI, Lang AS, Suttle CA**. 2007. The complete genomes of three viruses assembled from shotgun libraries of marine RNA virus communities. Virol J **4**:69. http://dx.doi.org/10.1186/1743-422X-4-69.

genomeA
Journals.ASM.org

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.

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 readthrough 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 and 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

1. **Bollback JP, Huelsenbeck JP.** 2001. Phylogeny, genome evolution, and host specificity of single-stranded RNA bacteriophage (family *Leviviridae*). J Mol Evol **52:**117–128.
2. **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.
3. **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.
4. **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.
5. **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.
6. **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.