genomeA
Journals.ASM.org

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.

Received 26 May 2015 **Accepted** 28 May 2015 **Published** 2 July 2015

Citation Greninger AL, DeRisi JL. 2015. Draft genome sequence of tombunodavirus UC1. Genome Announc 3(4):e00655-15. doi:10.1128/genomeA.00655-15.

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

Tombusviridae 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.

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. **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.

genomeA
Journals.ASM.org

Draft Genome Sequence of Picalivirus D 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 sequence of picalivirus D, a member of picalivirus, a picorna-like superfamily, and likely member of the order *Picornavirales*, assembled from metagenomic sequencing of organisms in San Francisco wastewater. This virus likely constitutes a novel genus within the picalivirus family.

Received 26 May 2015 **Accepted** 28 May 2015 **Published** 2 July 2015

Citation Greninger AL, DeRisi JL. 2015. Draft genome sequence of picalivirus D recovered from San Francisco wastewater. Genome Announc 3(4):e00654-15. doi:10.1128/genomeA.00654-15.

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). Picalivirus is an unassigned member within the picorna-like superfamily that consists of three viral genomes that were originally described in a worldwide survey of sewage enriched for viral particles (4). While originally described as having sequence motif hallmarks of both *Picornavirales* and *Caliciviridae*, picaliviruses increasingly appear to be members of the *Picornavirales* order.

While performing weekly metagenomic sequencing of organisms in San Francisco wastewater, we assembled a 9,570-nucleotide contig, which we are naming picalivirus D. This contig aligned by BLASTx to all three members of picalivirus. Picalivirus D consisted of two open reading frames (ORFs) of 7,023 and 2,238 nucleotides containing the nonstructural and structural proteins, respectively. This genome is 604 nucleotides longer than the complete genome of picalivirus A, which was attained by 5′ and 3′ rapid amplification of cDNA ends (RACE) (4). The RdRp portion of the first polyprotein of picalivirus D demonstrated 31 to 35% amino acid identity to picaliviruses A to C, with 27% amino acid identity to the predicted ABC ATPase/helicase of picalivirus A (the only other picalivirus with a sequence from that region). The capsid polyprotein aligned with 27 to 30% amino acid identity across the entirety of the polyprotein to picaliviruses A to C. The remaining highly ranking BLASTx hits to the picalivirus D genome all derived from the *Picornavirales* order, such as Taura syndrome virus, chicken picornavirus 1, *Aurantiochytrium* single-stranded RNA virus 01, and cricket paralysis virus.

The viral genome was 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 (5, 6). The average coverage of the three contigs using all reads from the sample was 294×.

The original sample was taken from wastewater on 25 January 2010, after a large rainstorm that left >5 inches of rain over the preceding week. 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. The 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. Other RNA viruses discovered in this sample include ciliovirus, brinovirus, laverivirus, tombunodavirus, and several other marine RNA viruses and phages (7–9).

Nucleotide sequence accession number. The GenBank accession number for picalivirus D is KF478837.

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. **Ng TF, Marine R, Wang C, Simmonds P, Kapusinszky B, Bodhidatta L, Oderinde BS, Wommack KE, Delwart E.** 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J Virol **86:**12161–12175. http://dx.doi.org/10.1128/JVI.00869-12.
5. **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.
6. **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.

7. **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. http://dx.doi.org/10.1128/genomeA.00653-15.
8. **Greninger AL, DeRisi JL.** 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.
9. **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. http://dx.doi.org/10.1128/genomeA.00652-15.

genomeA
Journals.ASM.org

Draft Genome Sequence of Laverivirus UC1, a Dicistrovirus-Like RNA Virus Featuring an Unusual Genome Organization

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 sequence of Laverivirus UC1, assembled from San Francisco wastewater. This dicistronic RNA virus bears some similarity to dicistroviruses; however, it appears to have a unique genome organization relative to all other known RNA viruses.

Received 25 May 2015 **Accepted** 28 May 2015 **Published** 2 July 2015

Citation Greninger AL, DeRisi JL. 2015. Draft genome sequence of Laverivirus UC1, a dicistrovirus-like RNA virus featuring an unusual genome organization. Genome Announc 3(4):e00656-15. doi:10.1128/genomeA.00656-15.

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

The picorna-like superfamily is a taxonomic unit of positive-stranded RNA viruses with conserved RNA-dependent RNA polymerase (RdRp), capsid, and helicase proteins that infect a broad range of hosts, including animals, plants, and insects (1–3). While performing weekly metagenomic sequencing of San Francisco wastewater, we assembled a contig of 8,440 nucleotides that yielded BLASTx alignment to the RNA-dependent RNA polymerase of the dicistrovirus *Solenopsis invicta virus-1* (30% amino acid identity), the ATPase/helicase of the dicistrovirus *Formica exsecta virus 1* (26% amino acid identity), and the capsid protein of the dicistrovirus aphid lethal paralysis virus (23% amino acid identity) (4, 5). The genome appears dicistronic encoding for two separate open reading frames (ORFs) in the standard genetic code of 2,094 and 5,604 nucleotides. However, unlike the dicistroviruses, the first ORF encodes for the RHV-like and CRPV-like capsid proteins, while the second ORF encodes for the ATPase/helicase and RNA-dependent RNA polymerase, similar to a dicistronic picorna-like superfamily organization. This differs from *Dicistroviridae* members that have the two ORFs in the opposite configuration (6). Noncoding regions include a 5′ untranslated region (UTR) of 493 nucleotides, an intergenic region of 162 nucleotides, and a 3′ UTR of 85 nucleotides.

We have given the name Laverivirus UC1 to this virus due to its discovery in wastewater and the Latin root "laver" for a water basin, as well as the lack of a known host. The host organisms of dicistroviruses are arthropods, though given the unique genome organization of Laverivirus this may not be the case. Of note, the plurality of nonchordate eukaryotic reads in the non-DNAsed sample aligned to the mountain pine beetle *Dendroctonus ponderosae* (7, 8). In the DNAsed sample, 3,512 reads were recovered to picornaviridae, including hits to human pathogens enterovirus, cardiovirus, parechovirus, Aichi virus, and salivirus (9–14).

This viral genome was recovered from a wastewater sample that was taken on 25 January 2010, one week after a large rainstorm. This same sample also contained novel ciliate and marine RNA viruses, along with a new picalivirus (15–18). 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-micron spin columns. The sample was treated with micrococcal nuclease, and nucleic acid was extracted using a Zymo Viral DNA/RNA kit; 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 (19, 20).

Nucleotide sequence accession number. The GenBank accession number for Laverivirus UC1 is KF510029.

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. **Van Munster M, Dullemans AM, Verbeek M, Van Den Heuvel JF, Clérivet A, Van Der Wilk F**. 2002. Sequence analysis and genomic organization of aphid lethal paralysis virus: a new member of the family *Dicistroviridae*. J Gen Virol **83:**3131–3138.
5. **Valles SM, Strong CA, Dang PM, Hunter WB, Pereira RM, Oi DH, Shapiro AM, Williams DF**. 2004. A picorna-like virus from the red imported fire ant, *Solenopsis invicta*: initial discovery, genome sequence, and characterization. Virology **328:**151–157. http://dx.doi.org/10.1016/ j.virol.2004.07.016.
6. **Bonning BC, Miller WA**. 2010. Dicistroviruses. Annu Rev Entomol **55:**129–150. http://dx.doi.org/10.1146/annurev-ento-112408-085457.
7. **Keeling CI, Yuen MM, Liao NY, Docking TR, Chan SK, Taylor GA, Palmquist DL, Jackman SD, Nguyen A, Li M, Henderson H, Janes JK,**

Zhao Y, Pandoh P, Moore R, Sperling FA, Huber DP, Birol I, Jones SJ, Bohlmann J. 2013. Draft genome of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major forest pest. Genome Biol **14:**R27. http://dx.doi.org/10.1186/gb-2013-14-3-r27.

8. **Mercado JE, Hofstetter RW, Reboletti DM, Negrón JF.** 2014. Phoretic symbionts of the mountain pine beetle (*Dendroctonus ponderosae* Hopkins). Forest Sci **60:**512–526.
9. **Yamashita T, Kobayashi S, Sakae K, Nakata S, Chiba S, Ishihara Y, Isomura S.** 1991. Isolation of cytopathic small round viruses with BS-C-1 cells from patients with gastroenteritis. J Infect Dis **164:**954–957. http://dx.doi.org/10.1093/infdis/164.5.954.
10. **Li L, Victoria J, Kapoor A, Blinkova O, Wang C, Babrzadeh F, Mason CJ, Pandey P, Triki H, Bahri O, Oderinde BS, Baba MM, Bukbuk DN, Besser JM, Bartkus JM, Delwart EL.** 2009. A novel picornavirus associated with gastroenteritis. J Virol **83:**12002–12006. http://dx.doi.org/10.1128/JVI.01241-09.
11. **Holtz LR, Finkbeiner SR, Zhao G, Kirkwood CD, Girones R, Pipas JM, Wang D.** 2009. Klassevirus 1, a previously undescribed member of the family *Picornaviridae*, is globally widespread. Virol J **6:**86. http://dx.doi.org/10.1186/1743-422X-6-86.
12. **Greninger AL, Runckel C, Chiu CY, Haggerty T, Parsonnet J, Ganem D, DeRisi JL.** 2009. The complete genome of klassevirus—a novel picornavirus in pediatric stool. Virol J **6:**82. http://dx.doi.org/10.1186/1743-422X-6-82.
13. **Greninger AL, Holtz L, Kang G, Ganem D, Wang D, DeRisi JL.** 2010. Serological evidence of human klassevirus infection. Clin Vaccine Immunol **17:**1584–1588. http://dx.doi.org/10.1128/CVI.00152-10.
14. **Chiu CY, Greninger AL, Kanada K, Kwok T, Fischer KF, Runckel C, Louie JK, Glaser CA, Yagi S, Schnurr DP, Haggerty TD, Parsonnet J, Ganem D, DeRisi JL.** 2008. Identification of cardioviruses related to Theiler's murine encephalomyelitis virus in human infections. Proc Natl Acad Sci U S A **105:**14124–14129. http://dx.doi.org/10.1073/pnas.0805968105.
15. **Greninger AL, DeRisi JL.** 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.
16. **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. http://dx.doi.org/10.1128/genomeA.00652-15.
17. **Greninger AL, DeRisi JL.** 2015. Draft genome 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.
18. **Ng TF, Marine R, Wang C, Simmonds P, Kapusinszky B, Bodhidatta L, Oderinde BS, Wommack KE, Delwart E.** 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J Virol **86:**12161–12175. http://dx.doi.org/10.1128/JVI.00869-12.
19. **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.
20. **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.