



Complete Genome Sequence of Pittsburgh Sewage-Associated Virus 1

Paul G. Cantalupo,^a James M. Pipas^a

^aDepartment of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

ABSTRACT We present the complete genome sequence of a virus found in raw sewage collected in Pittsburgh, PA, USA. Pittsburgh sewage-associated virus 1 (PSAV1) encodes one large open reading frame with conserved domains typically found in the *Picornavirales* order of viruses. PSAV1 is closely related to *Biomphalaria* virus 2 (BV2).

The *Picornavirales* order contains several virus families, such as *Picornaviridae*, *Dicistroviridae*, and *Secoviridae*, that infect mammals, insects, and plants. These viruses are small and nonenveloped and contain a monomeric single-strand positive-sense RNA genome of 7 to 10 kb in length. Usually they contain one or two large open reading frames (ORFs), which generate a polyprotein that is subsequently cleaved by a viral protease into functional polypeptides. The global diversity of these and other RNA viruses continues to be revealed through deep sequencing studies (1, 2). Raw sewage is a rich biomatrix that supports both multicellular and microorganism populations and has been shown to contain an abundance of novel viruses (3–5).

Viral nucleic acid from Pittsburgh raw sewage was extracted and sequenced (3). After nontemplate sequences were removed from the raw reads (NCBI Sequence Read Archive accession number SRR315458), reads were assembled with Celera assembler (v. 7.0) using a 3% utgErrorRate setting (6). The longest contig generated, 8,745 bp, was annotated in 2012 by a BLASTX (7) search, which showed it to be distantly related to *Dicistroviridae*. The termini of the genome were completed with 5' and 3' rapid amplification of cDNA ends (RACE), which added 49 and 201 bp to the 5' and 3' ends, respectively. The 3' RACE revealed that the genome is poly(A) tailed. The 8,995-bp genome sequence was confirmed by Sanger sequencing of staggered PCR amplicons across the genome.

A BLASTN search showed 78% identity between PSAV1 and *Biomphalaria* virus 2 (BV2) (GenBank accession number KY024322), an unclassified *Picornavirales* virus (8). PSAV1 contains one ORF that is predicted to encode a 2,668-amino acid (aa) polyprotein which is 86% identical to the BV2 polyprotein. PSAV1 polyprotein contains several functional domains, as identified by using CD-Search (9), including RNA helicase, protease, RNA-dependent RNA polymerase (RdRp), and three capsid domains. Since the proteolytic cleavage sites are unknown, the putative start and end coordinates of each polypeptide were based on the CD-Search results. The RNA-helicase (107 aa, 2,502 to 2822 nucleotides [nt]) contains a GX₄GKS motif and a KX₅SX₅S/TNN motif, which have been shown to be responsible for deoxynucleoside triphosphate (dNTP) binding and ATP hydrolysis (10). PSAV1 protease (196 aa, 3,558 to 4,145 nt) was putatively identified as a cysteine protease due to a cysteine at residue 1079 replacing the serine residue found in many RNA virus proteases (11). The RdRp (474 aa, 4,248 to 5,669 nt) polypeptide contains motifs A (DX₄D), B (PXG), and C (GDD), which are characteristic of a picornavirus RdRp (12). We were not able to identify a VPg protein in PSAV1. The three viral capsid polypeptides are predicted as capsid1 (195 aa, 6,318 to 6,902 nt), capsid2 (185 aa, 7,299 to 7,853 nt), and capsid3 (249 aa, 7,959 to 8,705 nt).

Received 28 November 2017 **Accepted** 23 February 2018 **Published** 15 March 2018

Citation Cantalupo PG, Pipas JM. 2018. Complete genome sequence of Pittsburgh sewage-associated virus 1. *Genome Announcements* 6:e01460-17. <https://doi.org/10.1128/genomeA.01460-17>.

Copyright © 2018 Cantalupo and Pipas. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to James M. Pipas, pipas@pitt.edu.

The phylogenetics of PSAV1 show that it is closely related to BV2, a snail virus. Since PSAV1 was isolated from raw sewage, its host is unknown. Both viruses share no significant homology to other viruses. These findings suggest that more members of this distant *Picornavirales* clade remain to be discovered.

Accession number(s). The genome sequence of PSAV1 has been deposited in GenBank under the accession number [MG552116](https://www.ncbi.nlm.nih.gov/nuccore/MG552116).

REFERENCES

1. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, Chen LJ, Qin XC, Xu J, Holmes EC, Zhang YZ. 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* 4:e05378. <https://doi.org/10.7554/eLife.05378>.
2. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539–543. <https://doi.org/10.1038/nature20167>.
3. Cantalupo PG, Calgua B, Zhao G, Hundesa A, Wier AD, Katz JP, Grabe M, Hendrix RW, Girones R, Wang D, Pipas JM. 2011. Raw sewage harbors diverse viral populations. *mBio* 2:e00180-11. <https://doi.org/10.1128/mBio.00180-11>.
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. <https://doi.org/10.1128/JVI.00869-12>.
5. Rosario K, Nilsson C, Lim YW, Ruan Y, Breitbart M. 2009. Metagenomic analysis of viruses in reclaimed water. *Environ Microbiol* 11:2806–2820. <https://doi.org/10.1111/j.1462-2920.2009.01964.x>.
6. Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G. 2008. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* 24:2818–2824. <https://doi.org/10.1093/bioinformatics/btn548>.
7. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
8. Galinier R, Tetreau G, Portet A, Pinaud S, Duval D, Gourbal B. 2017. First characterization of viruses from freshwater snails of the genus *Biomphalaria*, the intermediate host of the parasite *Schistosoma mansoni*. *Acta Trop* 167:196–203. <https://doi.org/10.1016/j.actatropica.2016.12.021>.
9. Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 32:W327–W331. <https://doi.org/10.1093/nar/gkh454>.
10. Smiley JR, Chang KO, Hayes J, Vinje J, Saif LJ. 2002. Characterization of an enteropathogenic bovine calicivirus representing a potentially new calicivirus genus. *J Virol* 76:10089–10098. <https://doi.org/10.1128/JVI.76.20.10089-10098.2002>.
11. 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 Micro* 6:925–939. <https://doi.org/10.1038/nrmicro2030>.
12. Hansen JL, Long AM, Schultz SC. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 5:1109–1122. [https://doi.org/10.1016/S0969-2126\(97\)00261-X](https://doi.org/10.1016/S0969-2126(97)00261-X).