



Near-Complete Genome Sequence of a Novel Single-Stranded RNA Virus Discovered in Indoor Air

 Karyna Rosario,^a  Noah Fierer,^{b,c}  Mya Breitbart^a

^aCollege of Marine Science, University of South Florida, Saint Petersburg, Florida, USA

^bDepartment of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado, USA

^cCooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA

ABSTRACT Viral metagenomic analysis of heating, ventilation, and air conditioning (HVAC) filters recovered the near-complete genome sequence of a novel virus, named HVAC-associated RNA virus 1 (HVAC-RV1). The HVAC-RV1 genome is most similar to those of picorna-like viruses identified in arthropods but encodes a small domain observed only in negative-sense single-stranded RNA viruses.

A novel RNA virus was detected through metagenomic sequencing of partially purified virus particles from heating, ventilation, and air conditioning (HVAC) filters used as passive indoor air samplers in university dormitory rooms (1). This virus, named HVAC-associated RNA virus 1 (HVAC-RV1), shared low sequence identity with known positive-sense single-stranded RNA (ssRNA) viruses. While HVAC-RV1 was originally identified in a male-occupied dorm room, metagenomic sequences with significant similarities to HVAC-RV1 were identified in 50% ($n = 12$) of the dormitory rooms sampled. The divergent HVAC-RV1 sequence, combined with its prevalence, prompted further investigation and confirmation of this novel viral genome.

A 9.8-kb contig sequence representing HVAC-RV1 was retrieved after assembling quality-filtered metagenomic reads (raw reads available under Sequence Read Archive [SRA] accession number SRR5853132) using SPAdes version 3.6 (2). Based on BLASTx comparisons against the GenBank nonredundant (nr) database, the contig sequence was most similar to Milolii virus, an RNA viral genome identified in ghost ants (3). However, the HVAC-RV1 contig sequence was smaller than the Milolii virus genome (~10.5 kb) and contained a major open reading frame (ORF) that was truncated (i.e., no stop codon). Therefore, the complete coding region of the assembled HVAC-RV1 genome sequence was verified through PCR and Sanger sequencing to eliminate potential assembly errors. In addition, the 3' end of the assembled sequence was extended using an oligo(dT) adapter primer provided with the rapid amplification of cDNA ends (RACE) system (Invitrogen) (4), indicating that the genome is polyadenylated.

The verified near-complete HVAC-RV1 genome sequence is ~10.6 kb in size and contains two major nonoverlapping ORFs flanked by 5' and 3' untranslated regions. Searches against the NCBI Conserved Domain Database (5) (E value < 0.001) revealed that ORF1, which is ~6.4 kb in length, contains RNA helicase (pfam00910), trypsin-like serine peptidase (pfam13365), and RNA-dependent RNA polymerase (pfam00680) conserved domains. This domain organization within polyproteins is reminiscent of picorna-like viruses (6, 7). However, unlike known picorna-like nonsegmented viral genomes, HVAC-RV1 encodes a second major ORF immediately downstream from ORF1 that contains both nonstructural and structural domains. ORF2, which is ~3.4 kb in length, contains domains with significant similarities to bunyavirus nonstructural NS-S (pfam03231) and calicivirus coat (pfam00915) proteins. The presence of a domain

Received 14 February 2018 **Accepted** 27 February 2018 **Published** 22 March 2018

Citation Rosario K, Fierer N, Breitbart M. 2018. Near-complete genome sequence of a novel single-stranded RNA virus discovered in indoor air. *Genome Announc* 6:e00198-18. <https://doi.org/10.1128/genomeA.00198-18>.

Copyright © 2018 Rosario et al. 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 Karyna Rosario, krosari2@mail.usf.edu.

similar to the bunyavirus NS-S protein family is surprising, since this domain has only been reported from negative-sense ssRNA viruses from the order *Bunyavirales* (8).

BLASTp comparisons of HVAC-RV1 ORF1 and ORF2 hypothetical protein sequences revealed limited similarities (< 40% identity) to polyproteins from the Milolii virus and other unclassified arthropod-associated viruses whose picorna-like genomes are predicted to encode a single polyprotein. Therefore, HVAC-RV1 contains a unique genome organization and further illustrates that current taxonomic classification schemes need to be reassessed to accommodate expanding RNA viral diversity (9).

Accession number(s). The near-complete genome sequence of the HVAC-associated RNA virus 1 has been deposited in GenBank under the accession number [MG775312](https://www.ncbi.nlm.nih.gov/nuccore/MG775312).

ACKNOWLEDGMENTS

We thank Unilever Industries and the National Science Foundation (grant DEB-1239976) for financial support.

REFERENCES

- Rosario K, Fierer N, Miller S, Luongo J, Breitbart M. 2018. Diversity of DNA and RNA viruses in indoor air as assessed via metagenomic sequencing. *Environ Sci Technol* 52:1014–1027. <https://doi.org/10.1021/acs.est.7b04203>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Brettell LE, Mordecai GJ, Pachori P, Martin SJ. 2017. Novel RNA virus genome discovered in ghost ants (*Tapinoma melanocephalum*) from Hawaii. *Genome Announc* 5:e00669-17. <https://doi.org/10.1128/genomeA.00669-17>.
- Scotto-Lavino E, Du G, Frohman MA. 2007. 3' End cDNA amplification using classic RACE. *Nat Protoc* 1:2742. <https://doi.org/10.1038/nprot.2006.481>.
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203. <https://doi.org/10.1093/nar/gkw1129>.
- Le Gall O, Christian P, Fauquet CM, King AMQ, Knowles NJ, Nakashima N, Stanway G, Gorbalenya AE. 2008. *Picornavirales*, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Arch Virol* 153:715–727. <https://doi.org/10.1007/s00705-008-0041-x>.
- 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. <https://doi.org/10.1038/nrmicro2030>.
- Eifan S, Schnettler E, Dietrich I, Kohl A, Blomström A-L. 2013. Non-structural proteins of arthropod-borne bunyaviruses: roles and functions. *Viruses* 5:2447–2468. <https://doi.org/10.3390/v5102447>.
- Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, Qin X-C, Li J, Cao J-P, Eden J-S, Buchmann J, Wang W, Xu J, Holmes EC, Zhang Y-Z. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539–543. <https://doi.org/10.1038/nature20167>.