



First Complete Genome Sequence of Human Coronavirus HKU1 from a Nonill Bat Guano Miner in Thailand

Yutthana Joyjinda,^a Apaporn Rodpan,^a Pongtorn Chartpituck,^b Krairerk Suthum,^b Suphaluk Yaemsakul,^b Thaniwan Cheun-Arom,^c Saowalak Bunprakob,^a Kevin J. Olival,^d Martha M. Stokes,^e Thiravat Hemachudha,^a Supaporn Wacharapluesadee^a

^aThai Red Cross Emerging Infectious Disease-Health Science Centre, World Health Organization Collaborating Centre for Research and Training on Viral Zoonoses, Chulalongkorn Hospital, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

^bThe 5th Office of Disease Prevention and Control, Ratchaburi, Thailand

^cDepartment of Biology, Faculty of Science, Ramkhamhaeng University, Bangkok, Thailand

^dEcoHealth Alliance, New York, New York, USA

^eCooperative Biological Engagement Program, Defense Threat Reduction Agency (DTRA), Washington, DC, USA

ABSTRACT Human coronavirus HKU1 (HCoV-HKU1) was first detected in a patient with viral pneumonia from Hong Kong in 2004. Here, we report the first complete genome sequence of HCoV-HKU1 from Thailand, obtained from a nonill person who worked in a bat cave. Phylogenetic tree analysis revealed it as a group B HCoV-HKU1.

There are four species of endemic human coronavirus (HCoV) currently recognized by the International Committee for the Taxonomy of Viruses, namely, HCoV-OC43, -229E, -NL63, and -HKU1, and two epidemic CoVs, including severe acute respiratory syndrome (SARS)-CoV and the Middle East respiratory syndrome (MERS)-CoV discovered in 2003 and 2012, respectively (1). Bats are believed to be the ancestral hosts of alpha- and beta-CoV, including SARS-CoV and MERS-CoV (2, 3). A group C betacoronavirus (MERS-related CoV) was detected from dry bat guano collected from a cave in Thailand where bat guano is sold for use as fertilizer (4). Nasopharyngeal swabs of bat guano miners ($n = 34$) in Ratchaburi Province, Thailand, were collected and sent to the Thai Red Cross-Emerging Infectious Diseases laboratory to test for CoVs using conventional PCR targeting the betacoronavirus RNA-dependent RNA polymerase (RdRp) gene (5). One sample was positive for the coronavirus HKU1 strain, and its whole genome was sequenced by using next-generation sequencing (NGS).

The nasopharyngeal swab found CoV positive by conventional PCR was subjected to metagenomic sequencing on the Illumina MiSeq platform. RNA was extracted using a QIAamp viral RNA mini kit (Qiagen, Germany), followed by DNase treatment. A DNA library was prepared using a TruSeq total RNA with Ribo-zero globin kit (Illumina). The quality and quantity of the DNA library was estimated by using the QIAxcel Advanced system and QIAseq library quantification kit, respectively (Qiagen, Germany). The 17 pmol of DNA library was injected into the flow cell and sequenced using a MiSeq reagent kit version 3.

A total of 59,704,490 sequencing reads from 2×150 -bp paired ends which, passing a quality score of 30, was used for data analysis. Host DNA (human genome from GATK resource bundle) removal was done using the SNAP (version 0.15.4) alignment tool (6). Mapping assembly was performed using HKU1 (GenBank accession number [AY884001](https://www.ncbi.nlm.nih.gov/nuccore/AY884001)) as the reference strain by using the Burrows-Wheeler Aligner (BWA) program (7). A total of 6,475,827 reads mapped to HCoV-HKU1 at a depth of 29,000 \times , which covered the whole genome of 29,811 bp. The consensus sequence of HCoV-HKU1 with a G+C

Citation Joyjinda Y, Rodpan A, Chartpituck P, Suthum K, Yaemsakul S, Cheun-Arom T, Bunprakob S, Olival KJ, Stokes MM, Hemachudha T, Wacharapluesadee S. 2019. First complete genome sequence of human coronavirus HKU1 from a nonill bat guano miner in Thailand. *Microbiol Resour Announc* 8:e01457-18. <https://doi.org/10.1128/MRA.01457-18>.

Editor Jason E. Stajich, University of California, Riverside

Copyright © 2019 Joyjinda 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 Supaporn Wacharapluesadee, spwa@hotmail.com.

Received 30 October 2018

Accepted 11 January 2019

Published 7 February 2019

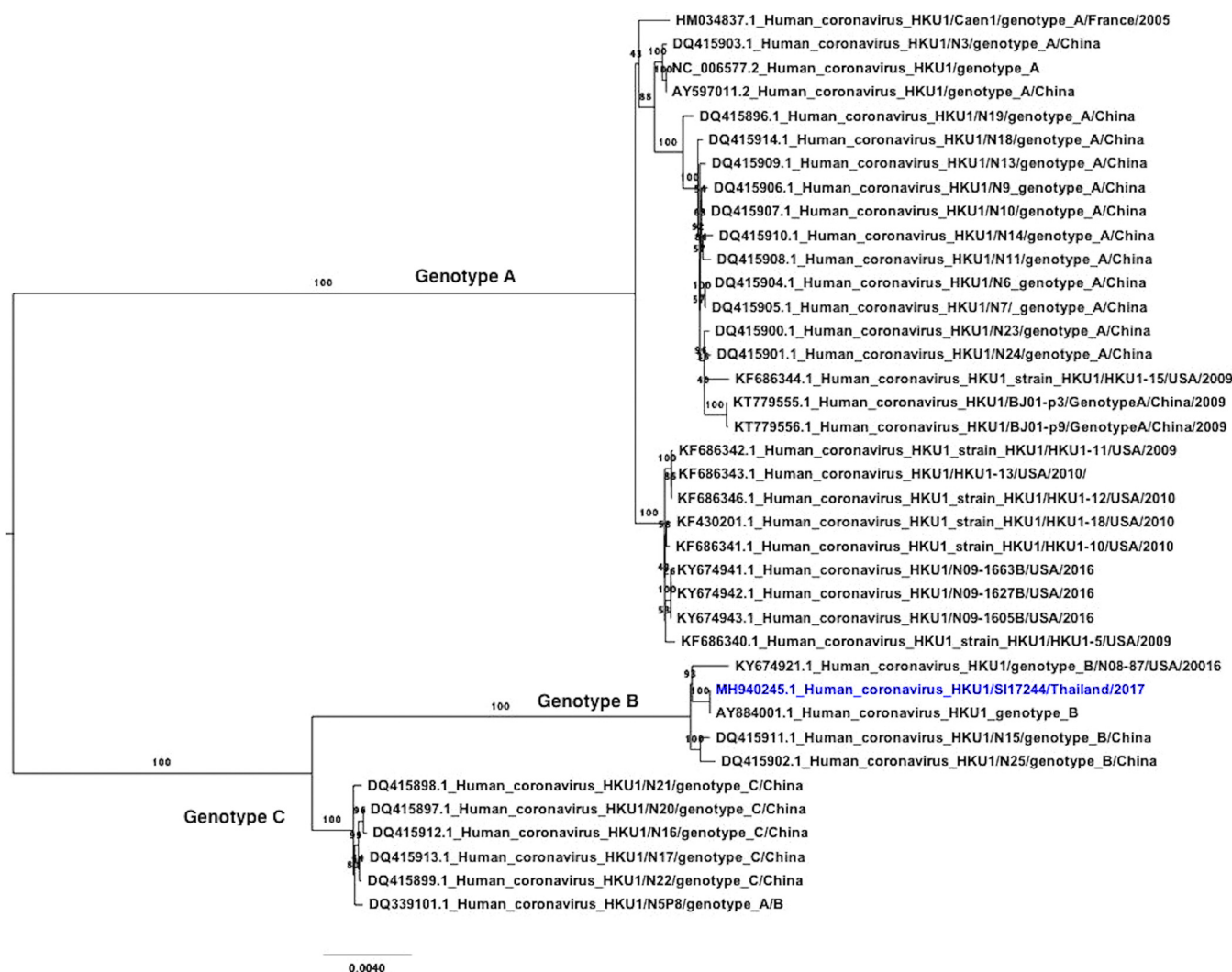


FIG 1 Phylogenetic tree of HCoV-HKU1 using maximum likelihood method with 1,000 bootstrap replicates. The whole-genome sequence of HCoV-HKU1 (29,811 bp) from Thailand is colored in blue; it belonged to group B genotype and was closely related to HCoV-HKU1 from Hong Kong (GenBank accession number [AY884001](https://www.ncbi.nlm.nih.gov/nuccore/AY884001)).

content of 65% was generated using SAMtools mpileup (8). This is the first complete genome of HCoV-HKU1 isolated from Thailand.

Phylogenetic analysis of our HCoV-HKU1 whole-genome and reference sequences from the GenBank database, using maximum likelihood algorithm on the MEGA6 program, showed that our sequence is in the same clade with sequences from China and the United States (Fig. 1). It belonged to the group B genotype and was closely related to HCoV-HKU1 from Hong Kong (GenBank accession number [AY884001](https://www.ncbi.nlm.nih.gov/nuccore/AY884001)) that was isolated in 2006 (99% nucleotide identity) (9). It is worth noting that this virus was detected from an individual with a high level of occupational exposure to bat feces and, thus, an elevated level of zoonotic virus spillover. While this individual was likely exposed to HCoV-HKU1 due to person-to-person transmission and not via exposure from bats, our surveillance strategy and viral characterization pipeline provide valuable insight into the circulation of endemic infectious diseases in Thailand and increase the country's preparedness for other novel emerging infectious diseases.

Data availability. The human coronavirus HKU1 strain reported here was deposited in GenBank under the accession number [MH940245](https://www.ncbi.nlm.nih.gov/nuccore/MH940245) and SRA accession number [PRJNA509533](https://www.ncbi.nlm.nih.gov/sra/PRJNA509533).

ACKNOWLEDGMENTS

Generous support for this work was provided by the Research Chair grant (P-13-01091); Cluster and Program Management Office (CPMO) (P-15-50535); the National Science and Technology Development Agency (NSTDA), Thailand; the American people through the United States Agency for International Development (USAID) Emerging Pandemic Threats PREDICT project (Cooperative Agreement number AID-OAA-A-14-00102), and The Defense Threat Reduction Agency Biological Threat Reduction Program (DTRA BTRP) (Cooperative Agreement number HDTRA1-17-C-0004 P00001).

REFERENCES

- Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. 2018. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res* 46:D708–D717. <https://doi.org/10.1093/nar/gkx932>.
- Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang YJ, Luo CM, Tan B, Wang N, Zhu Y, Crameri G, Zhang SY, Wang LF, Daszak P, Shi ZL. 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503:535–538. <https://doi.org/10.1038/nature12711>.
- Anthony SJ, Gilardi K, Menachery VD, Goldstein T, Ssebide B, Mbabazi R, Navarrete-Macias I, Liang E, Wells H, Hicks A, Petrosov A, Byarugaba DK, Debbink K, Dinnon KH, Scobey T, Randell SH, Yount BL, Cranfield M, Johnson CK, Baric RS, Lipkin WI, Mazet JA. 2017. Further evidence for bats as the evolutionary source of Middle East respiratory syndrome coronavirus. *mBio* 8:e00373-17. <https://doi.org/10.1128/mBio.00373-17>.
- Wacharapluesadee S, Sintunawa C, Kaewpom T, Khongnomnan K, Olival KJ, Epstein JH, Rodpan A, Sangsri P, Intarut N, Chindamporn A, Suksawa K, Hemachudha T. 2013. Group C betacoronavirus in bat guano fertilizer, Thailand. *Emerg Infect Dis* Aug 19:1349–1351. <https://doi.org/10.3201/eid1908.130119>.
- Watanabe S, Masangkay JS, Nagata N, Morikawa S, Mizutani T, Fukushi S, Alviola P, Omatsu T, Ueda N, Iha K, Taniguchi S, Fujii H, Tsuda S, Endoh M, Kato K, Tohya Y, Kyuwa S, Yoshikawa Y, Akashi H. 2010. Bat coronaviruses and experimental infection of bats, the Philippines. *Emerg Infect Dis* 16:1217–1223. <https://doi.org/10.3201/eid1608.100208>.
- Zaharia M, Bolosky WJ, Curtis K, Fox A, Patterson D, Shenker S, Stoica I, Karp RM, Sittler T. 2011. Faster and more accurate sequence alignment with SNAP. arXiv:1111.5572 [q-bio.GN].
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Woo PCY, Lau SKP, Yip CCY, Huang Y, Tsoi HW, Chan KH, Yuen KY. 2006. Comparative analysis of 22 Coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in Coronavirus HKU1. *J Virol* 80:7136–7145. <https://doi.org/10.1128/JVI.00509-06>.