



Complete Genome Sequences of Five Human Coronavirus NL63 Strains Causing Respiratory Illness in Hospitalized Children in China

Lu Zhang,^a Mian Gan,^b Zhaoyong Zhang,^b Xin Li,^b Wenkuan Liu,^b Airu Zhu,^b Jing Sun,^b Fang Li,^b Yanqun Wang,^b Fuchun Zhang,^a Jingxian Zhao,^b Rong Zhou,^b Jincun Zhao^{a,b}

^aInstitute of Infectious Disease, Guangzhou Eighth People's Hospital of Guangzhou Medical University, Guangzhou, China

^bState Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

ABSTRACT We report the complete genome sequences of five human coronavirus NL63 (HCoV-NL63) strains obtained using next-generation sequencing. The five HCoV-NL63 strains were obtained from hospitalized children with severe acute respiratory infection detected in Guangdong, China. This study provides several complete genomes of HCoV-NL63 and improves our understanding of HCoV-NL63 evolution in China.

Human coronavirus NL63 (HCoV-NL63) is a member of the family *Coronaviridae*, genus *Alphacoronavirus*, and was first discovered in 2004 (1). HCoV-NL63 is mainly associated with the common cold in children, the elderly, and immunocompromised patients (2, 3). The genome of HCoV-NL63 is about 27 kb with a conserved gene order of 5'-orf1ab-spike (S)-orf3-envelope (E)-membrane (M)-nucleocapsid (N)-poly (A). The species tropism of HCoV-NL63 is determined by spike glycoprotein. HCoV-NL63 and severe acute respiratory syndrome coronavirus (SARS-CoV) share the same cell receptor, angiotensin converting enzyme 2 (ACE-2) (4, 5), for entry into host cells, and HCoV-NL63 is recognized as a common cause of upper respiratory tract infection and has been prevalent worldwide.

Here, nasopharyngeal swab samples were collected from hospitalized children with severe acute respiratory infection in Guangzhou, China, in 2018. This study was performed in strict accordance with human subject protection guidance provided by the Research Ethics Committee of Guangzhou Medical University. The respiratory samples were filtered with 0.22- μ m filters, RNA extraction was performed using a Qiagen viral RNA extraction kit, and extracted RNA was used for sequence-independent single-primer amplification (SISPA) (6, 7) as follows: a reverse transcription reaction was performed with SuperScript III reverse transcriptase using a primer containing a fixed sequence, followed by a random hexamer at the 3' end (FR26RV, GCCGGAGCTCTGCAGATATC NNNNNN). Then, Klenow fragment polymerase (New England Biolabs) was used for DNA synthesis. Finally, PCR amplification was conducted using primers consisting of the fixed portions of the random primers (FR26, GCCGGAGCTCTGCAGATATC). Purified DNA was used for next-generation sequencing (NGS). Libraries were prepared with the Nextera XT kit (Illumina), and paired-end reads (2×125 bp) determined using a HiSeq 2500 instrument were used for cleaning and assembling using CLC Genomics Workbench version 11.0. Illumina sequencing yielded about 10 million reads per sample. Reads were assembled into contigs with a *de novo* assembly model, and the contig sequences were then extracted for subsequent analysis. Partial genome sequences of five HCoV-NL63 strains were obtained by NGS methods. Meanwhile, sets of specific primer pairs were

Citation Zhang L, Gan M, Zhang Z, Li X, Liu W, Zhu A, Sun J, Li F, Wang Y, Zhang F, Zhao J, Zhou R, Zhao J. 2020. Complete genome sequences of five human coronavirus NL63 strains causing respiratory illness in hospitalized children in China. *Microbiol Resour Announc* 9:e01597-19. <https://doi.org/10.1128/MRA.01597-19>.

Editor Simon Roux, DOE Joint Genome Institute

Copyright © 2020 Zhang 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 Rong Zhou, zhourong@vip.163.com, or Jincun Zhao, zhaojincun@gird.cn.

Received 17 January 2020

Accepted 4 February 2020

Published 20 February 2020

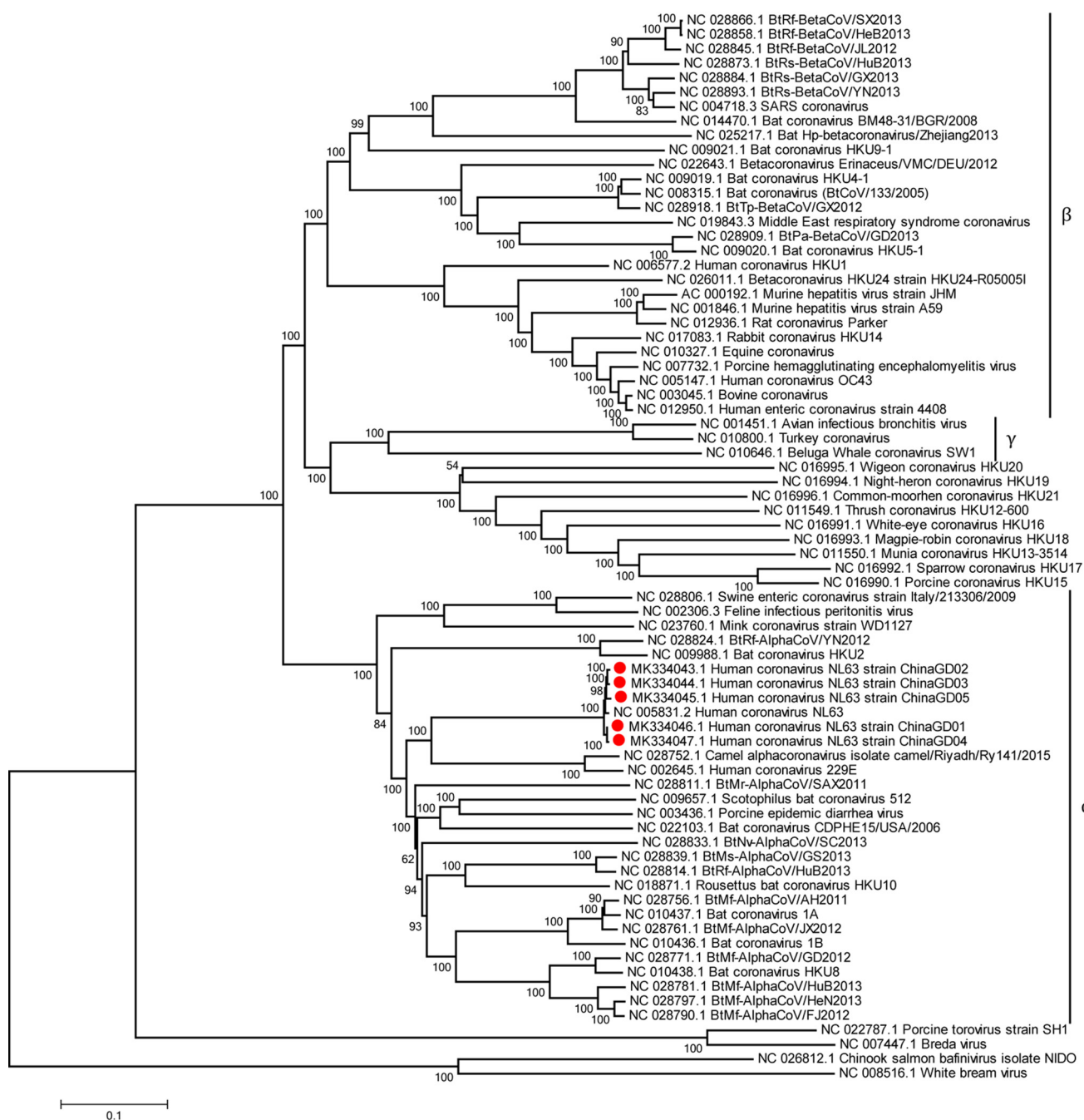
TABLE 1 Primers used for the genome sequencing of HCoV-NL63

Primer	Sequence (5'–3')	Target (nucleotide position) or amplification method	Size (bp)
1F	CCTGGCCTCTTGCTTTTTCACATGT	20504	1,686
1R	ACTTCGACGGTTGAGAAACAATAG	22189	1,686
2F	CGCGTTAAGAGTGGTTCACCAGGTG	22059	1,623
2R	CAAAGCTGCAAGCCGTCAGTAATT	23681	1,623
3F	TTCAATCAAGCCGATCAACAAGTT	23624	1,600
3R	GTCATCAATTAATCGAAGGAACATC	25223	1,600
4F	CGAAGAGCCTGTTGTTGGTATAGTC	25153	1,690
4R	AACACGCTTCCAACGAGGTTTCTTC	26849	1,690
5F	CCAGGGCTGATAAGCCTTCTCAGTT	26800	754
5R	GTGTATCCATATCAAAAACAATATC	27553	754
6F	TGAGGATGTTTGTGTTTGTGTTGAC	18864	1,710
6R	GTCAGGAACACCTAATTGTAACATA	20573	1,710
7F	TGCGTGGTTGGTTGGTATGGATGT	17234	1,682
7R	ACGCTCATAACGAACCTGAATACTA	18915	1,682
8F	ATTCAGCAACTGGTTCCTTAGATGT	15479	1,808
8R	GTTATCGCCACAAACATGAGCACTT	17286	1,808
9F	CTCCCTACTATGACACAGCTGAATC	14005	1,596
9R	AGCCGCAAAGAGTCTAAGTGTATCT	15600	1,596
10F	GACCGTACAACCTATTCAAAGTGTG	12398	1,659
10R	GTTCTTTACCACTAATAGCATACTT	14056	1,659
11F	GGGCTATGGCTAATGTTATACAAG	9801	1,435
11R	TTTGCATATTCATGGCACGCTTCA	11235	1,435
12F	ACCCTTCAGAGTGTGCTTCATCAT	11090	1,378
12R	AGTCGAGCTGCACTAGAACCCTTG	12467	1,378
13F	CAACCACTGTAAGTCTTCATGG	7758	2,103
13R	CTGCCAAAATAGAATAGCACTCAAC	9860	2,103
14F	GTCAAAGGGTGATGCTGAAGAGGC	5424	2,394
14R	TCAACTGACCAATCTCAATGTACTT	7817	2,394
15F	TAGAGATGAATTGGGTGTTCTGTGTT	3424	2,043
15R	GGTCCAACATCACCTGTAACAAATT	5466	2,043
16F	GCAGATGTTCCAGATGCTTTTCAAT	1637	1,864
16R	GCAACTGTACAAGTGTGGTACTAAT	3500	1,864
17F	CAGCAATTATGTTCTTCAGGACTTT	565	1,118
17R	GTGTAATGTGCGATAAACTGATTG	1682	1,118
18F	CTTAAAGAATTTTCTATCTATAGA	1	1,056
18R	CATGCACCAACACTCCAACCTCTCAG	1056	1,056
GSP 1	CGAAGAGCCTGTTGTTGGTATAGTC	3' RACE ^a	Unknown
AP	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT	3' RACE	Unknown
GSP 2	CCAGGGCTGATAAGCCTTCTCAGTT	Nested PCR	Unknown
AUAP	GGCCACGCGTCGACTAGTAC	Nested PCR	Unknown
AAP	GGCCACGCGTCGACTAGTACGGGGGGGGGG	5' RACE	Unknown
GSP 1	GTGTAATGTGCGATAAACTGATTG	5' RACE	Unknown
AUAP	GGCCACGCGTCGACTAGTAC	Nested PCR	Unknown
GSP 2	CATGCACCAACACTCCAACCTCTCAG	Nested PCR	Unknown
GSP3	CCATGGCCAAAAACAACATCAAAGT	Nested PCR	Unknown

^a RACE, rapid amplification of cDNA ends.

designed and used to amplify the gap region of HCoV-NL63, which was used for genome assemblies using the SeqMan subprogram of the DNASTar software version 7.1.0 with default parameters (Table 1). Finally, five complete genome sequences of HCoV-NL63 were obtained using next-generation sequencing and Sanger sequencing methods together and were designated strains ChinaGD01 (27,531 bp), ChinaGD02 (27,516 bp), ChinaGD03 (27,516 bp), ChinaGD04 (27,532 bp), and ChinaGD05 (27,544 bp). The five HCoV-NL63 strains presented here were aligned using MAFFT version 7.158 (8) and showed 98.5 to ~99.1% nucleotide homology with the prototype HCoV-NL63 virus (GenBank accession number [NC_005831.2](https://www.ncbi.nlm.nih.gov/nuccore/NC_005831.2)) as estimated using MEGA version 5.10 software (9) (Fig. 1).

Only two complete genome sequences of HCoV-NL63 associated with acute respiratory illness have been obtained and reported in China. The complete genome sequence data from our study will provide insight into the evolution and genetic diversity of HCoV-NL63 in China.



Downloaded from <http://mra.asm.org/> on April 17, 2020 by guest

FIG 1 Phylogenetic analysis of HCoV-NL63 based on complete genomes. The complete genomes of 69 coronavirus references were obtained from the GenBank database, and multiple alignments were performed using MAFFT version 7.158 with default parameters. The phylogenetic tree was constructed using the neighbor-joining method with 1,000 bootstraps in MEGA version 5.10 with default parameters. The numbers at the nodes represent bootstrap support. Bootstrap values greater than 70% were considered statistically significant for grouping. The HCoV-NL63 strains presented in this study are marked with red circles.

Data availability. The complete genome sequences of the five newly identified HCoV-NL63 strains have been deposited in GenBank under the accession numbers [MK334043](#), [MK334044](#), [MK334045](#), [MK334046](#), and [MK334047](#). The sequencing reads are available in the SRA database under BioProject accession number [PRJNA601331](#).

ACKNOWLEDGMENTS

This research was supported by grants from the National Key Research and Development Program of China (2018YFC1200100), National Natural Science Foundation of

China (NSFC 81702047, 81772191, 91842106, and 8181101118), State Key Laboratory of Respiratory Disease (SKLRD-QN-201715 and SKLRD-QN-201912), National Key Technology R&D Program (2018YFC1311900), Guangdong Science and Technology Foundation (2019B030316028), 2015 Thousand Talents Plan Award of China, China Postdoctoral Science Foundation, and the Ph.D. Start-up Fund of the Natural Science Foundation of Guangdong Province, China.

REFERENCES

1. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, Wertheim-van Dillen PM, Kaandorp J, Spaargaren J, Berkhout B. 2004. Identification of a new human coronavirus. *Nat Med* 10: 368–373. <https://doi.org/10.1038/nm1024>.
2. van der Hoek L, Pyrc K, Berkhout B. 2006. Human coronavirus NL63, a new respiratory virus. *FEMS Microbiol Rev* 30:760–773. <https://doi.org/10.1111/j.1574-6976.2006.00032.x>.
3. Kaiser L, Regamey N, Roiha H, Deffernez C, Frey U. 2005. Human coronavirus NL63 associated with lower respiratory tract symptoms in early life. *Pediatr Infect Dis J* 24:1015–1017. <https://doi.org/10.1097/01.inf.0000183773.80217.12>.
4. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pöhlmann S. 2005. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc Natl Acad Sci U S A* 102:7988–7993. <https://doi.org/10.1073/pnas.0409465102>.
5. Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, Huan Y, Yang P, Zhang Y, Deng W, Bao L, Zhang B, Liu G, Wang Z, Chappell M, Liu Y, Zheng D, Leibbrandt A, Wada T, Slutsky AS, Liu D, Qin C, Jiang C, Penninger JM. 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med* 11:875–879. <https://doi.org/10.1038/nm1267>.
6. Reyes GR, Kim JP. 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Mol Cell Probes* 5:473–481. [https://doi.org/10.1016/s0890-8508\(05\)80020-9](https://doi.org/10.1016/s0890-8508(05)80020-9).
7. Wang Y, Zhu N, Li Y, Lu R, Wang H, Liu G, Zou X, Xie Z, Tan W. 2016. Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China. *Clin Microbiol Infect* 22:458.e451–458.e459. <https://doi.org/10.1016/j.cmi.2016.01.006>.
8. Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>.
9. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739. <https://doi.org/10.1093/molbev/msr121>.