

A single nucleotide distinguishes the SARS-CoV-2 in the Wuhan outbreak in December 2019 from that in Beijing-Xinfadi in June 2020, China

Y.-Q. Pan¹, A. N. Bahoussi¹, F. Guo¹ and L. Xing^{1,2,3}

1) Institute of Biomedical Sciences, 2) Key Laboratory of Medical Molecular Cell Biology of Shanxi Province, Shanxi University and 3) Shanxi Provincial Key Laboratory for Prevention and Treatment of Major Infectious Diseases, Taiyuan, China

Abstract

Two major locally transmitted outbreaks of coronavirus disease 2019 occurred in China, one in Wuhan from December 2019 to April 2020, another in Beijing-Xinfadi in June 2020. Severe acute respiratory syndrome coronavirus 2 isolated from these two outbreaks can be distinguished by a conserved pyrimidine nucleotide located at nucleotide position 241 in the 5'-untranslated region of the virus genome.

© 2020 The Authors. Published by Elsevier Ltd.

Keywords: 5'-untranslated region, Beijing, coronavirus disease 2019, pyrimidine, Wuhan

Original Submission: 1 September 2020; **Revised Submission:** 8 December 2020; **Accepted:** 10 December 2020

Article published online: 15 December 2020

Corresponding author: L. Xing, Institute of Biomedical Sciences, Shanxi University, 92 Wucheng Road, Taiyuan, 030006, Shanxi Province, China.

E-mail: xingli107@gmail.com

Y.-Q. Pan, A.N. Bahoussi and F. Guo contributed equally to this article.

Coronavirus disease 2019 (COVID-19), an ongoing pandemic, is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. Like SARS-CoV-1, which caused the 2002–2004 SARS outbreak [2], SARS-CoV-2 is also a member of the genus β -coronavirus [3]. The genome of SARS-CoV-2 is a positive-sense, single-stranded RNA that is ~29.8 kb in length and possesses 14 open reading frames [4].

COVID-19 was first identified in December 2019 in Wuhan, Hubei province of China [5,6]. The rapid local transmission of SARS-CoV-2 resulted in the lockdown of Wuhan city from 23 January until 8 April 2020. On 11 June 2020, after 55 days without a locally transmitted COVID-19 case, another metropolitan area, Beijing, the capital of China, reported the first COVID-19 case [7]. By 19 June 2020, 205 cases were officially confirmed and most of these patients went to Beijing-Xinfadi wholesale market. This second outbreak of locally transmitted

COVID-19 in Beijing-Xinfadi led to a partial lockdown of Beijing.

On 17 June 2020, three complete genomic sequences of the SARS-CoV-2 isolated in Beijing-Xinfadi were available on the GISAID initiative database (<https://www.gisaid.org>). We compared these newly identified viral sequences with that isolated from the first locally transmitted COVID-19 outbreak in Wuhan and found a uniquely conserved pyrimidine nucleotide at nucleotide position (nt) 241 relative to the SARS-CoV-2 reference sequence (NCBI reference no. NC_045512.2) (Fig. 1a). This conserved pyrimidine nucleotide is located in the loop region of stem-loop 5B (SL5B) [8] of the 5'-untranslated region of the virus genome (Fig. 1b) and can be used to distinguish the virus in Beijing-Xinfadi from that in Wuhan. Viruses isolated from both the environment and the patients in Beijing-Xinfadi have a uridine nucleotide, whereas the viruses isolated in Wuhan have a cytidine nucleotide in the corresponding genomic position (Fig. 1a). Moreover, from December 2019 to March 2020, a total of 210 complete viral genome sequences isolated in Wuhan and deposited in the GISAID database all had a cytidine nucleotide at nt 241 (data not shown). Interestingly, five coronaviruses isolated from pangolin (*Manis javanica*) in 2017 have a uridine nucleotide at nt 241, which is the same as SARS-CoV-2 in Beijing-Xinfadi (Fig. 1a). In contrast, the virus isolated from bat (*Rhinolophus affinis*) in 2013 (hCoV-19/bat/Yunnan/RaTG13/2013) has a cytidine nucleotide

a

Location	Host	Virus name	Accession ID	Sampling date	Sequence (5' to 3')	Nucleotide position
Guangxi	Pangolin	hCoV-19/pangolin/Guangxi/P4L/2017	EPI_ISL_410538	2017	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	269
	Pangolin	hCoV-19/pangolin/Guangxi/P1E/2017	EPI_ISL_410539	2017	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	269
	Pangolin	hCoV-19/pangolin/Guangxi/P5L/2017	EPI_ISL_410540	2017	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	269
	Pangolin	hCoV-19/pangolin/Guangxi/P5E/2017	EPI_ISL_410541	2017	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	269
	Pangolin	hCoV-19/pangolin/Guangxi/P2V/2017	EPI_ISL_410542	2017	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	251
Beijing (Xinfadi)	Environment	hCoV-19/env/Beijing/IVDC-03-06/2020	EPI_ISL_469256	2020-06-11	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	261
	Human	hCoV-19/Beijing/IVDC-02-06/2020	EPI_ISL_469255	2020-06-11	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	262
	Human	hCoV-19/Beijing/IVDC-01-06/2020	EPI_ISL_469254	2020-06-11	UUUUGUCCGGGUGUGACC GAAAGGUAAGAUG	268
Yunnan	Bat	hCoV-19/bat/Yunnan/RaTG13/2013	EPI_ISL_402131	2013-07-24	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	253
	Environment	hCoV-19/env/Wuhan/IVDC-HBF13-20/2020	EPI_ISL_408514	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	268
	Environment	hCoV-19/env/Wuhan/IVDC-HBF13-21/2020	EPI_ISL_408515	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	268
	Human	hCoV-19/Wuhan/IVDC-HB-GX02/2019	EPI_ISL_434534	2019-12-30	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	268
	Human	hCoV-19/Wuhan/HB-CDC-HB-03/2019	EPI_ISL_412899	2019-12-30	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	257
Hubei (Wuhan)	Human	hCoV-19/Wuhan/HB-CDC-HB-02/2019	EPI_ISL_412898	2019-12-30	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	257
	Environment	hCoV-19/env/Wuhan/IVDC-HBF54/2020	EPI_ISL_408512	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	144
	Human	hCoV-19/Wuhan/WH03/2020	EPI_ISL_406800	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	248
	Human	hCoV-19/Wuhan/WHU02/2020	EPI_ISL_406717	2020-01-02	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	267
	Human	hCoV-19/Wuhan/WHU01/2020	EPI_ISL_406716	2020-01-02	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	267
	Human	hCoV-19/Wuhan/IPBCAMS-WH-05/2020	EPI_ISL_403928	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	268
	Human	hCoV-19/Wuhan/IVDC-HB-04/2020	EPI_ISL_402120	2020-01-01	UUUCGUCCGGGUGUGACC GAAAGGUAAGAUG	268

*** *****

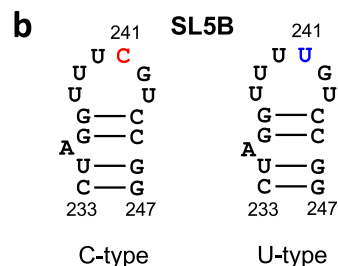


FIG. 1. (a) Alignment of genomic RNA sequence of indicated animal and human β -coronaviruses. The information was obtained from the GISAID EpiCoV™ database. Alignment was performed using EMBL-EBI Clustal Omega. Numbers indicating nucleotide position are relative to the genomic sequence of the indicated viruses. (b) The partial sequence and secondary structure of stem-loop 5B (SL5B) in the 5'-untranslated region of SARS-CoV-2. The conservation of both sequence and secondary structure of viral RNA was analysed in combination using the online LocARNA program [11,12] (<http://rna.informatik.uni-freiburg.de/>). Numbers indicating nucleotide position are relative to the SARS-CoV-2 reference sequence (NCBI reference no. NC_045112.2). The conserved pyrimidine nucleotide is shown with red or blue letters.

(Fig. 1a). It appears that the mutation of this conserved pyrimidine nucleotide at nt 241 in the 5'-untranslated region of the viral genome occurred before the outbreak of COVID-19. This single pyrimidine nucleotide can serve as a biomarker to distinguish two major locally transmitted outbreaks of COVID-19 in China.

Phylogenetic analysis of virus evolution is currently used to infer the route of infection transmission, which is based, with bias, on the coding region of the viral genome [4,9]. The non-coding 5'-untranslated region of coronavirus directly interacts with the cellular factors to regulate transcription, translation and viral genome synthesis [10]. The character of this non-coding region is an additive indicator for virus evolution. Therefore, the identification of this pyrimidine nucleotide has implications for inferring the transmission route of new COVID-19 cases in China.

Funding

This study was supported by the Shanxi Provincial Department of Science and Technology.

Conflict of interest

None declared.

Authors' contributions

LX conceived the study. YQP, ANB and FG collected and analysed the data. LX wrote the paper.

Acknowledgements

We thank all the persons who contributed to the production and collection of the virus sequences in GISAID database.

References

- [1] Coronaviridae Study Group of the International Committee on Taxonomy of V. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 2020;5:536–44.
- [2] Sorensen MD, Sorensen B, Gonzalez-Dosal R, Melchjorsen CJ, Weibel J, Wang J, et al. Severe acute respiratory syndrome (SARS): development of diagnostics and antivirals. *Ann N Y Acad Sci* 2006;1067:500–5.
- [3] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;579:270–3.
- [4] Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 2020;27:325–8.
- [5] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. China Novel Coronavirus I, Research T. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;382:727–33.
- [6] Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet* 2020;395:470–3.
- [7] Lew L. Beijing Covid-19 outbreak puts food markets back in infection focus. *South China Morning Post*; 2020.
- [8] Rangan R, Zheludev IN, Hagey RJ, Pham EA, Wayment-Steele HK, Glenn JS, et al. RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *RNA* 2020;26:937–59.
- [9] Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci U S A* 2020;117:9241–3.
- [10] Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* 2015;1282:1–23.
- [11] Raden M, Ali SM, Alkhnbashi OS, Busch A, Costa F, Davis JA, et al. Freiburg RNA tools: a central online resource for RNA-focused research and teaching. *Nucleic Acids Res* 2018;46:W25–9.
- [12] Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA* 2012;18:900–14.