Current Biology, Volume *30*

Supplemental Information

A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein

Hong Zhou, Xing Chen, Tao Hu, Juan Li, Hao Song, Yanran Liu, Peihan Wang, Di Liu, Jing Yang, Edward C. Holmes, Alice C. Hughes, Yuhai Bi, and Weifeng Shi

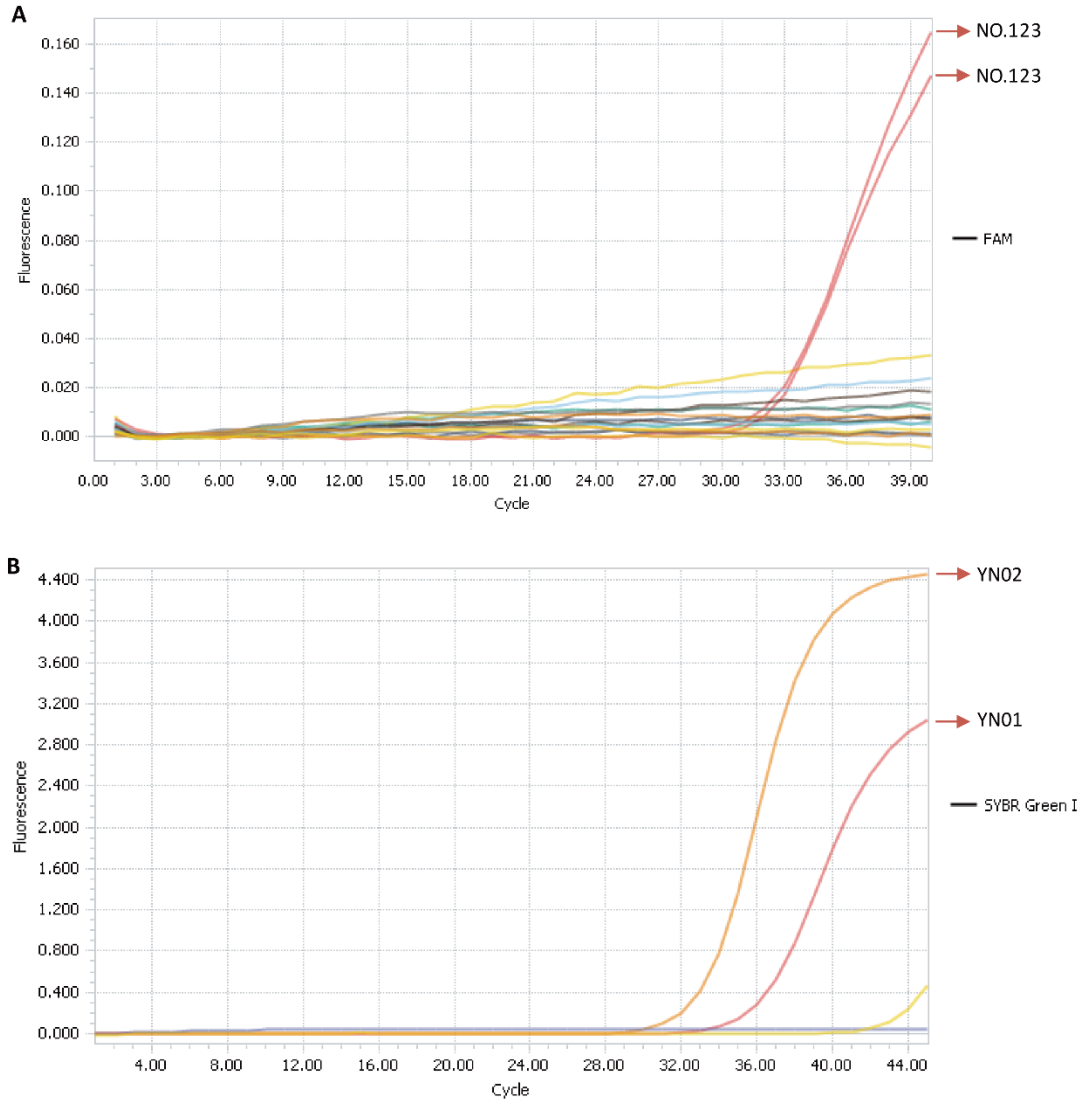

Figure S1. Detection of RmYN02 using real-time PCR primers and Taqman probes in the eight fecal samples from pool 39, related to Sanger sequencing of STAR Methods. Two replicates were set for each original sample. (A) TaqMan-based qPCR was performed to detect the existence of RmYN02 in the eight additional fece samples of pool 39. Two replicates were set for each original sample. (B) RmYN01 and RmYN02 were identified in sample 123 using Tip green supermix (TransGen). Nuclease-free water was used as negative control.

Figure S2. Comparison of NGS consensus and Sanger sequencing of RmYN02, related to Figure 2 and Genome assembly and annotation of STAR Methods. (A) NGS consensus and Sanger sequencing results of the RBD and the cleavage site of RmYN02. The insertion of the multiple amino acids at the S1/S2 cleavage site is highlighted. (B) Sites in the RmYN02 genome that display nucleotide polymorphisms in the NGS data. The positions of these sites in RmYN02 were provided at the bottom of the figure. "/": Sanger sequencing was not performed.

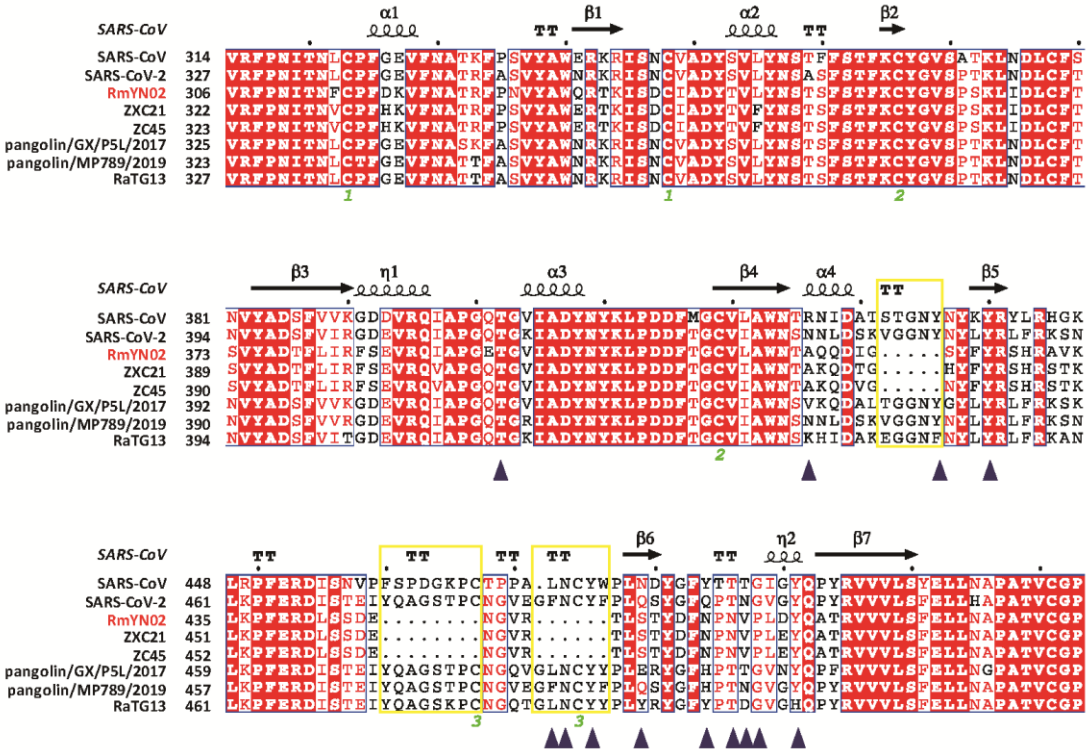

Figure S3. Sequence alignment of the RBDs from RmYN02 and representative beta-CoVs, related to Figure 2. The blue triangles indicate amino acids from the SARS-CoV S protein that impact binding to ACE2. The yellow rectangles highlight the three deletions in sequence resulting two loops shorten in RmYN02.

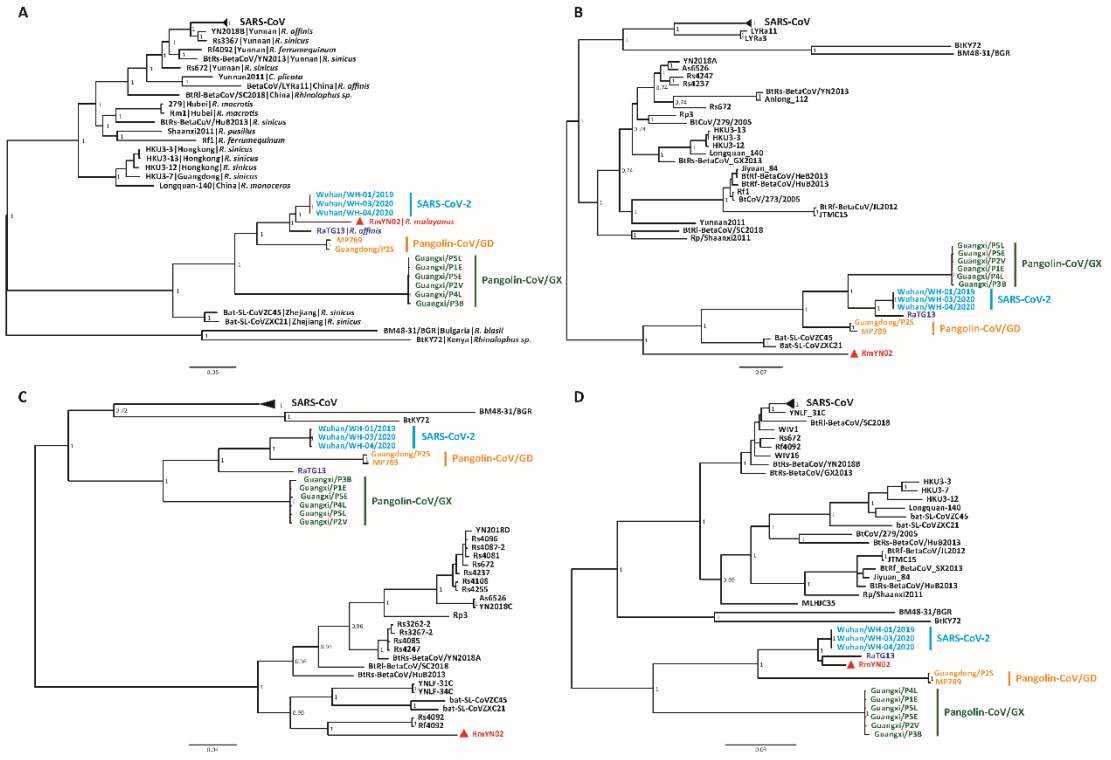

Figure S4. Bayesian phylogenetic analysis of SARS-CoV-2 and representative viruses of the subgenus *Sarbecoronavirus*, related to Figure 3. (A) Phylogenetic tree of the full-length virus genome. (B) the spike gene. (C) the RBD. (D) the RdRp gene. Phylogenetic analysis was also performed using MrBayes, employing the GTR nucleotide substitution model. Ten million steps were run, with trees and parameters sampled every 1,000 steps. The tree is midpoint rooted for clarity. RBD is delimited as the gene region 991-1572 of the spike gene according to the reference [6].

Species	Individual number							Number of Libraries	Sample number				
	May	Jun	Jul	Aug	Sep	Oct	May-Oct		Samples per library	Patagium	Lung	Liver	Feces
Rhinolophus malayanus[a]	6	24	9				39	5	10, 10, 8, 9, 11	37			11
Rhinolophus stheno[a]	14	12	16	6			48	7	10, 5, 10, 10, 9, 7, 9	44			16
Hipposideros larvatus (complex) [a]	4	8	12	5	7	5	41	8	10, 10, 10, 10, 7, 7, 1, 1	40	1	1	14
Rhinolophus sinicus[a]	4	4	4	3	2		17	3	10, 7, 8	17			8
Myotis laniger[a]	7		2				9	1	9,	9			
Rhinolophus siamensis[a]	1	3	1	2		1	8	2	8, 3	8			3
Hipposideros pomona[a]	1		4	7	1		13	3	8, 5, 8	13			8
Kerivoula hardwickii[a]	1			1	1		3	1	3,	3			
Murina cyclotis[a]	1	2	2	2			7	2	7, 3	7			3
Aselliscus stoliczkanus[a]	1			2	1	1	5	2	5, 3	5			3
Myotis muricola[a]		3	1	2	1	1	8	1	8,	8			
Kerivoula sp. [a]	2						2	2	2, 1	2			1
Rhinolophus paradoxolophus [a]	1						1	2	1, 1	1			1
Kerivoula papillosa[a]			1	1			2	1	2,	2			
Tylonycteris robustula[a]		1					1	1	1,	1			

Harpiocephalus harpia[a]				1			1	2	1, 1	1			1
Hipposideros armiger[a]					1	2	3	3	2, 1, 1	3			1
Rhinolophus pearsonii[a]				3		3	6	3	6, 2, 3	6			5
Chaerephon plicata[b]			4				4	3	4, 2, 1	4		1	2
Taphozous melanopogon[b]			9				9	4	8, 1, 1, 1	8	1	1	1
Total	43	57	65	35	14	13		56		219	2	3	78

Table S1. Summary of the bat samples collected in the present study and the pooling strategy used for next-generation sequencing, related to Sample collection and Next generation sequencing of STAR Methods. Note: [a]Samples were collected from Mengla, Xishuangbanna, Yunan (101.27156323E, 21.91889683N). [b]Samples were collected from Mengla, Xishuangbanna, Yunan (21.5932019N, 101.2200914E). Different colors represent different sample types. Patagium, black; Lung, purple; Liver, blue; Feces, red.

Primers (nucleotide positions)[a]	Sequence, 5'→3'	Product length, bp
Forward qF (21344-21366)	ACCCAATTCAGTTGTCTTCCTAT	146
Reverse qR (21469-21489)	TCTAACGATGAGCCTACCCTT	
Probe qP (21411-21438)	TGCTGTTATGTCTCTTAAGGAGGGACAA	
Forward qF (YN02-21711)	TAATTCCACTCTAACCTGGTATA	107
Reverse qR (YN02-21795)	CAGCATTTGATTTATCTACAGTT	
Forward qF (YN01-21311)	GCACCTCACTCAGGATTAT	89
Reverse qR (YN01-21381)	AATAGACTTGACGATCCGA	
Forward F6 (23070-23092)	TGGTGTTCTAACTGATTCAGATA	2370
Reverse R6 (25418-25439)	TCTCTTTTTAAGGGTTATGATT	
Forward F1 (12002-12024)	CTTTCCATGCAGGGTGCTGTAGA	1488
Reverse R1 (13470-13489)	CGCACGGTGTAAGACGGGCT	
Forward F2 (13159-13180)	TGGTCAGGCAATAACAGTTACA	2579
Reverse R2 (15715-15737)	GATGCGTAAGTGCTATTAAAACA	
Forward F3 (15625-15646)	AGAGATGTTGACACAGACTTTG	1788
Reverse R3 (17390-17412)	GTACACATAGTGCTTAGCACGTA	
Forward F4 (17321-17345)	CAGATATAGTTGTCTTTGATGAAAT	1937
Reverse R4 (19236-19257)	AGGCAAGTTAAGGTTAGATAGC	

Table S2. Oligonucleotide primers designed to detect RmYN02 and to amplify the spike and 1b genes, related to Sanger sequencing of STAR Methods. [a]Values in parentheses indicate primer positions corresponding to the genome sequence of RmYN02.

Accession ID	Virus name	Location	Collection date	Originating lab	Submitting lab	Authors
EPI_ISL_402131	BetaCoV/bat/Yunnan/RaTG13/2013	Asia / China / Yunnan / Pu'er	2013-07-24	Wuhan Institute of Virology, Chinese Academy of Sciences	Wuhan Institute of Virology, Chinese Academy of Sciences	Yan Zhu, Ping Yu, Bei Li, Ben Hu, Hao-Rui Si, Xing-Lou Yang, Peng Zhou, Zheng-Li Shi
EPI_ISL_410539	BetaCoV/ pangolin/ Guangxi/P1E/ 2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410541	BetaCoV/ pangolin/ Guangxi/P5E/ 2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410540	BetaCoV/ pangolin/ Guangxi/P5L/2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410538	BetaCoV/ pangolin/ Guangxi/P4L/2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang

EPI_ISL_410543	BetaCoV/pangolin/Guangxi/P3B/2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410542	BetaCoV/pangolin/Guangxi/P2V/2017	Asia / China / Guangxi	2017	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410544	BetaCoV/pangolin/Guangdong/P2S/2019	Asia / China / Guangdong	2019	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
MT084071.1	MP789	China	2019-03-29		Chinese Academy of Fishery Sciences (SCSFRI, CAFS)	Jiang, J.-Z., Liu, P. and Chen, J.-P.

Table S3. Acknowledgement of sharing of SARS-CoV-2 genome sequences from the GISAID and GenBank databases, related to Figure1. We gratefully thank the authors listed below for sharing their genomic sequences of coronaviruses analyzed in this study.