

Supplementary Materials

Supplementary Table S1. Previous studies on the possible roles of pangolins in the evolution and cross-species transmission of SARS-CoV-2 and SC2r-Cov

Methodology	Reference(s)	Natural host	Incidental host	Intermediate host
Genetic analysis:				
● The genomic divergence between pangolin-CoVs and SARS-CoV-2 is larger than that between bat-CoVs and SARS-CoV-2.	Lam, et al. 2020; Liu, et al. 2020; Zhang, et al. 2020	In favor	No conflict	Against
● The sequences of S protein RBDs from some pangolin-CoVs share the highest amino acid identity with SARS-CoV-2.	Xiao, et al. 2020	No conflict	No conflict	In favor
● Synonymous substitutions in the RBD suggest that RaTG13 is more closely related to SARS-CoV-2 than are pangolin-CoVs.	Lam, et al. 2020	No conflict	No conflict	Against
● There is high genetic diversity among the lineages of pangolin-CoVs.	Lam, et al. 2020; Liu, et al. 2020; Zhang, et al. 2020; This study	In favor	Against/No conflict*	No conflict
Clinical observation:				
● Some pangolins infected by pangolin-CoVs show disease symptoms.	Liu, et al. 2019; Liu, et al. 2020; Xiao, et al. 2020	Against	In favor	In favor
● The S protein constructs based on the sequence of pangolin-CoV bind the pangolin ACE2 receptor.	Niu, et al. 2021; Wrobel, et al. 2021	No conflict	In favor	In favor
● PCR screening detected positive signals for pangolin-CoVs in three pangolins seized in Guangdong without pathological symptoms.	Li, Wang, Hua, et al. 2021	In favor	No conflict	No conflict
Survey:				
● PCR screening of 334 live pangolins confiscated from Peninsular Malaysia and Borneo from 2009–2019 yielded no positive results.	Lee, et al. 2020	Against	In favor	No conflict
● The SARS-CoV-2 surrogate virus neutralization test detected one positive pangolin sample confiscated within 2020 in southern Thailand.	Wacharapluesadee, et al. 2021	In favor	No conflict	No conflict
● PCR screening and mtDNA analysis indicated that pangolins with positive of pangolin-CoVs are widely distributed in Southeast Asia.	This study	In favor	Against/No conflict*	No conflict

Note: *Multiple spillover events were required.

Supplementary Table S2. The representative 21 SC2r-CoV and SARSr-CoV genomes published previously

Host	Sample ID	Accession No.	Database	Reference
<i>Manis javanica</i>	GX/P5L	EPI_ISL_410540	GISAID	Lam, et al. 2020
<i>Rhinolophus malayanus</i>	RmYN02	EPI_ISL_412977	GISAID	Zhou, Chen, et al. 2020
<i>Manis javanica</i>	MP789	MT121216	GenBank	Liu, et al. 2020
<i>Rhinolophus affinis</i>	RaTG13	MN996532	GenBank	Zhou, Yang, et al. 2020
<i>Homo sapiens</i>	Wuhan-Hu-1	MN908947	GenBank	Wu, et al. 2020
<i>Rhinolophus pusillus</i>	ZC45	MG772933	GenBank	Hu, et al. 2018
<i>Rhinolophus pusillus</i>	ZXC21	MG772934	GenBank	Hu, et al. 2018
<i>Rhinolophus cornutus</i>	Rc-o319	LC556375	GenBank	Murakami, et al. 2020
<i>Rhinolophus shameli</i>	RshSTT200	EPI_ISL_852605	GISAID	Hul, et al. 2021
<i>Rhinolophus shameli</i>	RshSTT182	EPI_ISL_852604	GISAID	Hul, et al. 2021
<i>Rhinolophus acuminatus</i>	RacCS203	MW251308	GenBank	Wacharapluesadee, et al. 2021
<i>Rhinolophus sinicus</i>	WIV1	KF367457	GenBank	Ge, et al. 2013
<i>Rhinolophus sinicus</i>	HKU3-1	DQ022305	GenBank	Lau, et al. 2005
<i>Rhinolophus blasii</i>	BM48-31	NC_014470	GenBank	Drexler, et al. 2010
<i>Rhinolophus affinis</i>	LYRa11	KF569996	GenBank	He, et al. 2014
<i>Homo sapiens</i>	BJ01	AY278488	GenBank	Qin, et al. 2003
<i>Rhinolophus monaceros</i>	Longquan-140	KF294457	GenBank	Lin, et al. 2017
<i>Rhinolophus ferrumequinum</i>	HeB2013	KJ473812	GenBank	Wu, et al. 2016
<i>Rhinolophus ferrumequinum</i>	Rfl	DQ412042	GenBank	Li, et al. 2005
<i>Rhinolophus sp.</i>	BtKY72	KY352407	GenBank	Tao and Tong 2019
<i>Rhinolophus blythi</i>	PrC31	MW703458	GenBank	Li, Wang, Ma, et al. 2021

Supplementary Table S3. The number of viral RNA sequencing reads mapped to the 21 available SC2r-CoV and SARSr-CoV genomes

Species	Sample	Isolation date	NGS sequencing data	Number of reads mapped to the 21 SC2r-CoV and SARSr-CoV genomes
<i>Manis javanica</i>	MJ18007	2018/12/06	16.6G	0
<i>Manis pentadactyla</i>	MP20	2017/08/29	17G	12
<i>Manis javanica</i>	MJ57	2017/08/29	17.8G	8
<i>Manis javanica</i>	MJ20	2017/08/29	17.5G	0
<i>Manis javanica</i>	MJ54	2017/08/29	18.4G	0

The aligned length of PE150 read to at least one SC2r-CoV or SARSr-CoV genome > 100 nucleotides.

Supplementary Table S4. The divergence between MP20 and the reference genome of pangolin-CoV-GXC (GXP5L)

	dS	dN	dN/dS	MP20 length (nucleotide)	GXP5L length (nucleotide)	Aligned length (nucleotide)
Genome-average	0.1801	0.0106	0.0588	22446	29184	22446
<i>ORF1ab</i>	0.1853	0.0106	0.0571	16851	21264	16851
<i>S</i>	0.2422	0.0089	0.0369	2514	3801	2514
<i>ORF3a</i>	0.1821	0.0192	0.1054	642	825	642
<i>E</i>	/	/	/	0	225	0
<i>M</i>	0.0798	0.0023	0.0288	567	666	567
<i>ORF6</i>	0.2614	0.0133	0.0511	183	183	183
<i>ORF7a</i>	0.1641	0.0203	0.1235	324	363	324
<i>ORF7b</i>	/	/	/	0	129	0
<i>ORF8</i>	/	/	/	0	363	0
<i>N</i>	0.0667	0.0104	0.1564	1251	1251	1251
<i>ORF10</i>	0	0.0114	NA	114	114	114

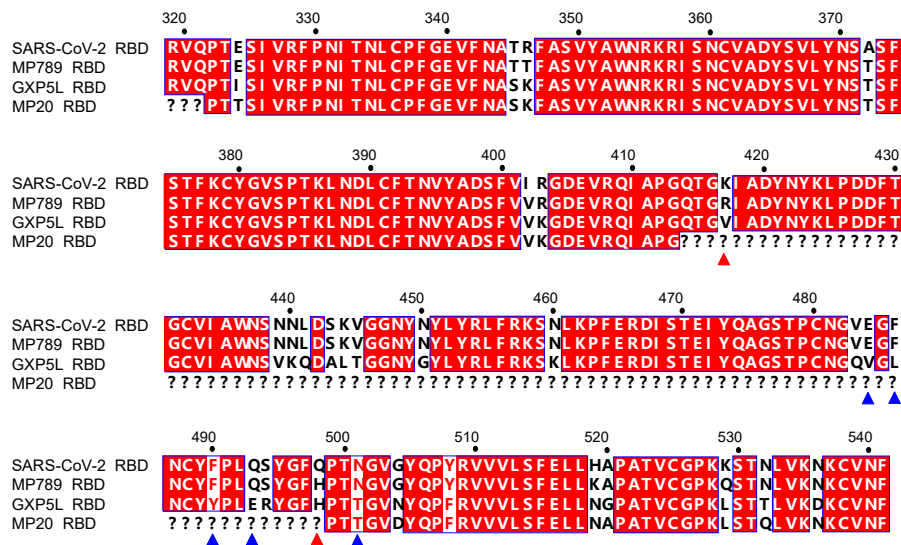
Supplementary Table S5. The divergence between pangolin-CoV-GDC (MP789) and pangolin-CoV-GXC (GXP5L)

	dS	dN	dN/dS	MP789 length (nucleotide)	GXP5L length (nucleotide)	Aligned length (nucleotide)
Genome-averag	0.715	0.060	0.084			
e	2	3	3	29139	29184	29097
ORF1ab	0.761	0.054	0.072			
	2	9	1	21267	21264	21237
S	0.956	0.099	0.104			
	2	7	2	3795	3801	3789
ORF3a	0.503	0.067	0.133			
	3	2	5	825	825	825
E	0.075	0.006	0.079			
	1		5	225	225	225
M	0.339	0.018	0.053			
	2		1	666	666	666
ORF6	0.316	0.041	0.130			
	8	3	4	183	183	183
ORF7a	0.518	0.076	0.147			
	4	5	6	363	363	363
ORF7b	0.492	0.171	0.347			
	6	1	3	129	129	129
ORF8	0.820	0.105	0.128			
	1	2	3	315	363	315
N	0.277	0.041	0.147			
	6		7	1257	1251	1251
ORF10	0.044	0.059	1.343			
	2	4	9	114	114	114

Supplementary Table S6. The number of putative iHVs identified in sample MP20

Position	Gene	Coverage	Reference (GXP5L)	Alternative allele	Annotation	Amino acid change
1299	<i>ORF1ab</i>	43	G (9)	T (34)	nonsynonymous	p.Gly345Cys
6105	<i>ORF1ab</i>	6	G (3)	T (3)	nonsynonymous	p.Leu1947Phe
8096	<i>ORF1ab</i>	6	T (3)	C (3)	synonymous	p.Thr2610Thr
25389	<i>ORF3</i>	7	A (4)	G (3)	synonymous	p.Val13Val
26530*	<i>M</i>	5	A (2)	C (2), T(1)	nonsynonymous	p.Leu17Phe
29624	3' UTR	5	C (2)	T (3)	intergenic_region	

*For position 26,530, both A>C and A>T cause Leu17Phe.



Supplementary Figure S1. Alignment of spike protein receptor-binding domain (RBD) sequences from SARS-CoV-2, pangolin-CoV-GDC (MP789), pangolin-CoV-GXC (GXP5L), and MP20.

The sites highlighted with colored triangles are the residue substitutions at the RBD binding sites of both pangolin-CoV-GDC and pangolin-CoV-GXC (two in red) and specifically of pangolin-CoV-GXC (five in blue) (Niu, et al. 2021). Identical residues are noted in red background, and similar residues are labeled in red boxed with blue frames. Two gaps (position 23266-23301 nt and 25577-25588 nt) filled by using MJ57 reads were not in the RBD (position 22488-23156 nt) of MP20.

REFERENCES

- Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, Seebens A, Niedrig M, Pfefferle S, Yordanov S, et al. 2010. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol* 84(21):11336-11349.
- Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, et al. 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503(7477):535-538.
- He B, Zhang Y, Xu L, Yang W, Yang F, Feng Y, Xia L, Zhou J, Zhen W, Feng Y, et al. 2014. Identification of diverse alphacoronaviruses and genomic characterization of a novel severe acute respiratory syndrome-like coronavirus from bats in China. *J Virol* 88(12):7070-7082.
- Hu D, Zhu C, Ai L, He T, Wang Y, Ye F, Yang L, Ding C, Zhu X, Lv R, et al. 2018. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg Microbes Infect* 7(1):154.
- Hul V, Delaune D, Karlsson EA, Hassanin A, Ou Tey P, Baidaliuk A, Gambaro F, Tu VT, Keatts L, Mazet J, et al. 2021. A novel SARS-CoV-2 related coronavirus in bats from Cambodia.

bioRxiv:10.1101/2021.1101.1126.428212.

- Lam TT, Jia N, Zhang YW, Shum MH, Jiang JF, Zhu HC, Tong YG, Shi YX, Ni XB, Liao YS, et al. 2020. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* 583(7815):282-285.
- Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH, Yuen KY. 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc Natl Acad Sci U S A* 102(39):14040-14045.
- Lee J, Hughes T, Lee MH, Field H, Rovie-Ryan JJ, Sitam FT, Sipangkui S, Nathan S, Ramirez D, Kumar SV, et al. 2020. No evidence of coronaviruses or other potentially zoonotic viruses in Sunda pangolins (*Manis javanica*) entering the wildlife trade via Malaysia. *Ecohealth* 17(3):406-418.
- Li L, Wang X, Hua Y, Liu P, Zhou J, Chen J, An F, Hou F, Huang W, Chen J. 2021. Epidemiological study of Betacoronaviruses in captive Malayan pangolins. *Front Microbiol* 12:657439.
- Li LL, Wang JL, Ma XH, Sun XM, Li JS, Yang XF, Shi WF, Duan ZJ. 2021. A novel SARS-CoV-2 related coronavirus with complex recombination isolated from bats in Yunnan province, China. *Emerg Microbes Infect* 10(1):1683-1690.
- Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, et al. 2005. Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310(5748):676-679.
- Lin XD, Wang W, Hao ZY, Wang ZX, Guo WP, Guan XQ, Wang MR, Wang HW, Zhou RH, Li MH, et al. 2017. Extensive diversity of coronaviruses in bats from China. *Virology* 507:1-10.
- Liu P, Chen W, Chen JP. 2019. Viral metagenomics revealed sendai virus and coronavirus infection of Malayan pangolins (*Manis javanica*). *Viruses* 11(11):979.
- Liu P, Jiang JZ, Wan XF, Hua Y, Li L, Zhou J, Wang X, Hou F, Chen J, Zou J, et al. 2020. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog* 16(5):e1008421.
- Murakami S, Kitamura T, Suzuki J, Sato R, Aoi T, Fujii M, Matsugo H, Kamiki H, Ishida H, Takenaka-Uema A, et al. 2020. Detection and characterization of bat sarbecovirus phylogenetically related to SARS-CoV-2, Japan. *Emerg Infect Dis* 26(12):3025-3029.
- Niu S, Wang J, Bai B, Wu L, Zheng A, Chen Q, Du P, Han P, Zhang Y, Jia Y, et al. 2021. Molecular basis of cross-species ACE2 interactions with SARS-CoV-2-like viruses of pangolin origin. *Embo j* 40(16):e107786.
- Qin E, Zhu Q, Yu M, Fan B, Chang G, Si B, Yang B, Peng W, Jiang T, Liu B, et al. 2003. A complete sequence and comparative analysis of a SARS-associated virus (Isolate BJ01). *Chin Sci Bull* 48(10):941-948.
- Tao Y, Tong S. 2019. Complete genome sequence of a severe acute respiratory syndrome-related coronavirus from Kenyan bats. *Microbiol Resour Announc* 8(28):e00548.
- Wacharapluesadee S, Tan CW, Maneeporn P, Duengkae P, Zhu F, Joyjinda Y, Kaewpom T, Chia WN, Ampoot W, Lim BL, et al. 2021. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat Commun* 12(1):972.
- Wrobel AG, Benton DJ, Xu P, Calder LJ, Borg A, Roustan C, Martin SR, Rosenthal PB, Skehel JJ, Gamblin SJ. 2021. Structure and binding properties of Pangolin-CoV spike glycoprotein inform the evolution of SARS-CoV-2. *Nat Commun* 12(1):837.
- Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, et al. 2020. A new coronavirus associated with human respiratory disease in China. *Nature* 579(7798):265-269.

- Wu Z, Yang L, Ren X, He G, Zhang J, Yang J, Qian Z, Dong J, Sun L, Zhu Y, et al. 2016. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *Isme j* 10(3):609-620.
- Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, Li N, Guo Y, Li X, Shen X, et al. 2020. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* 583(7815):286-289.
- Zhang T, Wu Q, Zhang Z. 2020. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol* 30(7):1346-1351.e1342.
- Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, Wang P, Liu D, Yang J, Holmes EC, et al. 2020. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr Biol* 30(11):2196-2203.e2193.
- Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, et al. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798):270-273.