SNV reoccurrence in multiple regions in the genome of 2019-nCoV

Xinjun He[1,*], Ph.D., Wei Yue[1], M.S., Jun Yan[1], Ph.D.

1 Affiliation:
Nanjing YiGenCloud Institute
211 Pubin Rd. Nanjing Jiangbei New Area, Nanjing, China, 210000

* **Correspondence:** xinjun.he@yigencloud.cn

Abstract:

The rapid development of 2019-2020 Wuhan seafood market pneumonia currently posed a major public health concern in China. Genome sequencing identified a novel beta-coronavirus closely related to SARS-CoV, named 2019-nCoV by WHO, as the cause of this pandemic disease. Viruses with single stranded RNA genome are prone to evolve quickly by accumulation of mutations, such as SNV, INDEL and cross viral recombination, aiding fast transmission among hosts and cross species. Here we collected related genome sequences and investigated variations shared by different strains of 2019-nCoV, identified reoccurrence of SNV mutations in clusters of patients, an indication of rapid evolution of 2019-nCoV at the transmission from animal host to human. The information collected herein would help to understand the dynamics of current pandemic.

Key words:

2019-nCoV, SARS-CoV-2, novel corona virus, Wuhan pneumonia

Introduction:

As of Feb. 2nd, 2020, the 2019-2020 Wuhan seafood market pneumonia pandemic has reached more than 17,000 infected cases and caused 361 deaths in China (information obtained from the website publication of the National Health Commission of China). It currently posed a major public health concern both in China and internationally with rapid development (*1-7*).

A quick search of the cause resulted in the identification of a novel coronavirus, named 2019-nCoV, which has not been previously reported, in the subgenus of sarbecovirus. DNA sequencing indicated that this coronavirus contained ~30kb of single strand RNA genome (ssRNA) (GenBank: MN908947, posting on Virological.org website, communicated by Edward C. Holmes, University of Sydney on behalf of the consortium led by Professor Yong-Zhen Zhang, Fudan University, Shanghai, http://virological.org/t/novel-2019-coronavirus-genome/319)(*8*). Multiple studies analyzed different strains of 2019-nCoV revealed its close relationship with SARS-CoV that caused 2002-2003 SARS pandemic pneumonia, and with several bat SARS-CoV related genomes, supporting the argument that 2019-nCoV may have transmitted from bat to human as well (*2, 6, 9*). In-depth phylogenic analysis indicated that 2019-nCoV shares 96% identity with a bat coronavirus strain RaTG13 and 79% of SARS-CoV (*6, 10*). It has been evidenced that frequent recombination happens among related coronavirus and fast evolution of 2019-nCoV genome with mutations has been proposed after among human transmission (*10-14*). Here we investigated publicly available 2019-nCoV genome sequences for such evidence.

Results:

We downloaded 53 genome sequences from 2019-nCoV and SARS, SARS-like virus sequences from GenBank, National Genomics Data Center of Beijing Institute of Genomics

(BIGD), and Global Initiative on Sharing All Influenza Data (GISAID) databases (**Table 1**). 53 complete genome sequences were then aligned with ClustalO algorithms online server at EBI (https://www.ebi.ac.uk/Tools/msa/clustalo/)(*15*). Alignments of sequences then were imported into JalViewer 2.0.11.0(*16*) for visualization. We also downloaded the metadata and mutation mapping data against 2019-nCoV reference genome (GenBank: NC_045512) from 2019 Novel Coronavirus Resource of BIGD to gain information of genome variations.

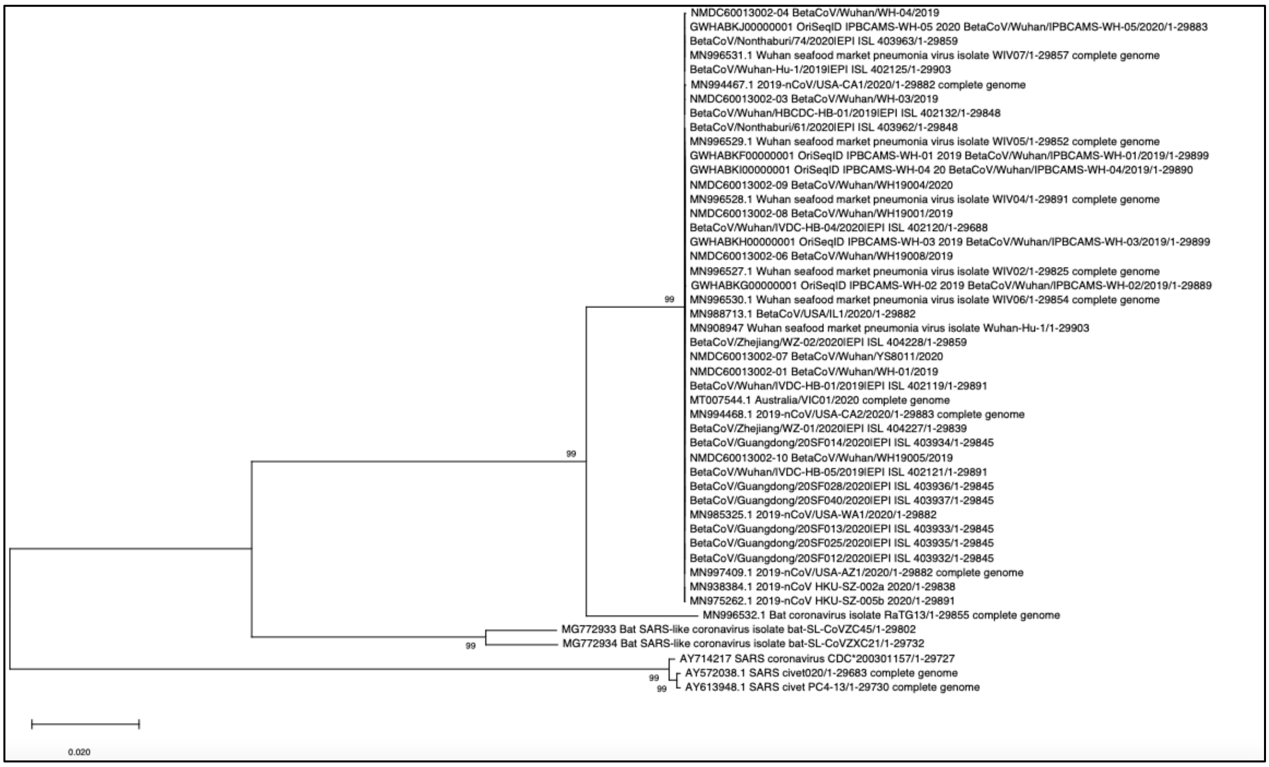
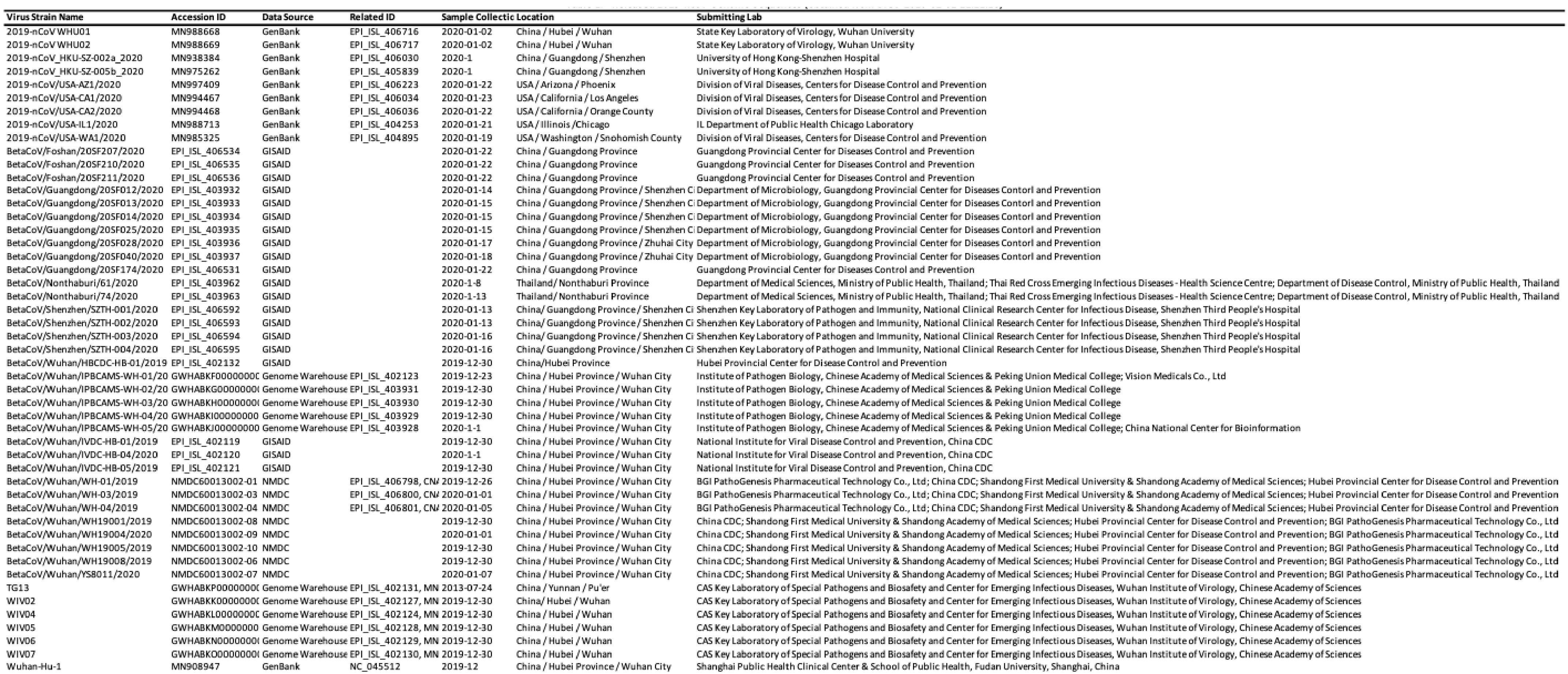


Figure 1. Phylogenetic tree reconstruction of 2019-nCoV with related SARS viral genomes. Whole genome sequences of 2019-nCoV, two bat SARS-like sequence, RaTG13, two civet SARS and one SARS-CoV from human were aligned and phylogenetic tree was reconstructed with NJ with 100 bootstrap replications. SARS sequences were used to root the tree.

Table 1. 2019-nCoV Genome Sequences (Obtained from BIGD 2020-02-02 for genome wide mutation analysis)*



Virus Strain Name	Accession ID	Data Source	Related ID	Sample Collectic	Location	Submitting Lab
2019-nCoV WHU01	MN988668	GenBank	EPI_ISL_406716	2020-01-02	China / Hubei / Wuhan	State Key Laboratory of Virology, Wuhan University
2019-nCoV WHU02	MN988669	GenBank	EPI_ISL_406717	2020-01-02	China / Hubei / Wuhan	State Key Laboratory of Virology, Wuhan University
2019-nCoV_HKU-SZ-002a_2020	MN938384	GenBank	EPI_ISL_406030	2020-1	China / Guangdong / Shenzhen	University of Hong Kong-Shenzhen Hospital
2019-nCoV_HKU-SZ-005b_2020	MN975262	GenBank	EPI_ISL_405839	2020-1	China / Guangdong / Shenzhen	University of Hong Kong-Shenzhen Hospital
2019-nCoV/USA-AZ1/2020	MN997409	GenBank	EPI_ISL_406223	2020-01-22	USA / Arizona / Phoenix	Division of Viral Diseases, Centers for Disease Control and Prevention
2019-nCoV/USA-CA1/2020	MN994467	GenBank	EPI_ISL_406034	2020-01-23	USA / California / Los Angeles	Division of Viral Diseases, Centers for Disease Control and Prevention
2019-nCoV/USA-CA2/2020	MN994468	GenBank	EPI_ISL_406036	2020-01-22	USA / California / Orange County	Division of Viral Diseases, Centers for Disease Control and Prevention
2019-nCoV/USA-IL1/2020	MN988713	GenBank	EPI_ISL_404253	2020-01-21	USA / Illinois /Chicago	IL Department of Public Health Chicago Laboratory
2019-nCoV/USA-WA1/2020	MN985325	GenBank	EPI_ISL_404895	2020-01-19	USA / Washington / Snohomish County	Division of Viral Diseases, Centers for Disease Control and Prevention
BetaCoV/Foshan/20SF207/2020	EPI_ISL_406534	GISAID		2020-01-22	China / Guangdong Province	Guangdong Provincial Center for Diseases Control and Prevention
BetaCoV/Foshan/20SF210/2020	EPI_ISL_406535	GISAID		2020-01-22	China / Guangdong Province	Guangdong Provincial Center for Diseases Control and Prevention
BetaCoV/Foshan/20SF211/2020	EPI_ISL_406536	GISAID		2020-01-22	China / Guangdong Province	Guangdong Provincial Center for Diseases Control and Prevention
BetaCoV/Guangdong/20SF012/2020	EPI_ISL_403932	GISAID		2020-01-14	China / Guangdong Province / Shenzhen Ci	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF013/2020	EPI_ISL_403933	GISAID		2020-01-15	China / Guangdong Province / Shenzhen Ci	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF014/2020	EPI_ISL_403934	GISAID		2020-01-15	China / Guangdong Province / Shenzhen Ci	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF025/2020	EPI_ISL_403935	GISAID		2020-01-15	China / Guangdong Province / Shenzhen Ci	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF028/2020	EPI_ISL_403936	GISAID		2020-01-17	China / Guangdong Province / Zhuhai City	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF040/2020	EPI_ISL_403937	GISAID		2020-01-18	China / Guangdong Province / Zhuhai City	Department of Microbiology, Guangdong Provincial Center for Diseases Contorl and Prevention
BetaCoV/Guangdong/20SF174/2020	EPI_ISL_406531	GISAID		2020-01-22	China / Guangdong Province	Guangdong Provincial Center for Diseases Control and Prevention
BetaCoV/Nonthaburi/61/2020	EPI_ISL_403962	GISAID		2020-1-8	Thailand/ Nonthaburi Province	Department of Medical Sciences, Ministry of Public Health, Thailand; Thai Red Cross Emerging Infectious Diseases - Health Science Centre; Department of Disease Control, Ministry of Public Health, Thailand
BetaCoV/Nonthaburi/74/2020	EPI_ISL_403963	GISAID		2020-1-13	Thailand/ Nonthaburi Province	Department of Medical Sciences, Ministry of Public Health, Thailand; Thai Red Cross Emerging Infectious Diseases - Health Science Centre; Department of Disease Control, Ministry of Public Health, Thailand
BetaCoV/Shenzhen/SZTH-001/2020	EPI_ISL_406592	GISAID		2020-01-13	China/ Guangdong Province / Shenzhen Ci	Shenzhen Key Laboratory of Pathogen and Immunity, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital
BetaCoV/Shenzhen/SZTH-002/2020	EPI_ISL_406593	GISAID		2020-01-13	China/ Guangdong Province / Shenzhen Ci	Shenzhen Key Laboratory of Pathogen and Immunity, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital
BetaCoV/Shenzhen/SZTH-003/2020	EPI_ISL_406594	GISAID		2020-01-16	China/ Guangdong Province / Shenzhen Ci	Shenzhen Key Laboratory of Pathogen and Immunity, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital
BetaCoV/Shenzhen/SZTH-004/2020	EPI_ISL_406595	GISAID		2020-01-16	China/ Guangdong Province / Shenzhen Ci	Shenzhen Key Laboratory of Pathogen and Immunity, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital
BetaCoV/Wuhan/HBCDC-HB-01/2019	EPI_ISL_402132	GISAID		2019-12-30	China/Hubei Province	Hubei Provincial Center for Disease Control and Prevention
BetaCoV/Wuhan/IPBCAMS-WH-01/20	GWHABKF0000000	Genome Warehouse	EPI_ISL_402123	2019-12-23	China / Hubei Province / Wuhan City	Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College; Vision Medicals Co., Ltd
BetaCoV/Wuhan/IPBCAMS-WH-02/20	GWHABKG0000000	Genome Warehouse	EPI_ISL_403931	2019-12-30	China / Hubei Province / Wuhan City	Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College
BetaCoV/Wuhan/IPBCAMS-WH-03/20	GWHABKH0000000	Genome Warehouse	EPI_ISL_403930	2019-12-30	China / Hubei Province / Wuhan City	Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College
BetaCoV/Wuhan/IPBCAMS-WH-04/20	GWHABKI00000000	Genome Warehouse	EPI_ISL_403929	2019-12-30	China / Hubei Province / Wuhan City	Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College
BetaCoV/Wuhan/IPBCAMS-WH-05/20	GWHABKJ00000000	Genome Warehouse	EPI_ISL_403928	2020-1-1	China / Hubei Province / Wuhan City	Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College; China National Center for Bioinformation
BetaCoV/Wuhan/IVDC-HB-01/2019	EPI_ISL_402119	GISAID		2019-12-30	China / Hubei Province / Wuhan City	National Institute for Viral Disease Control and Prevention, China CDC
BetaCoV/Wuhan/IVDC-HB-04/2020	EPI_ISL_402120	GISAID		2020-1-1	China / Hubei Province / Wuhan City	National Institute for Viral Disease Control and Prevention, China CDC
BetaCoV/Wuhan/IVDC-HB-05/2019	EPI_ISL_402121	GISAID		2019-12-30	China / Hubei Province / Wuhan City	National Institute for Viral Disease Control and Prevention, China CDC
BetaCoV/Wuhan/WH-01/2019	NMDC60013002-01	NMDC	EPI_ISL_406798, CN	2019-12-26	China / Hubei Province / Wuhan City	BGI PathoGenesis Pharmaceutical Technology Co., Ltd; China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention
BetaCoV/Wuhan/WH-03/2019	NMDC60013002-03	NMDC	EPI_ISL_406800, CN	2020-01-01	China / Hubei Province / Wuhan City	BGI PathoGenesis Pharmaceutical Technology Co., Ltd; China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention
BetaCoV/Wuhan/WH-04/2019	NMDC60013002-04	NMDC	EPI_ISL_406801, CN	2020-01-05	China / Hubei Province / Wuhan City	BGI PathoGenesis Pharmaceutical Technology Co., Ltd; China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention
BetaCoV/Wuhan/WH19001/2019	NMDC60013002-08	NMDC		2019-12-30	China / Hubei Province / Wuhan City	China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention; BGI PathoGenesis Pharmaceutical Technology Co., Ltd
BetaCoV/Wuhan/WH19004/2020	NMDC60013002-09	NMDC		2020-01-01	China / Hubei Province / Wuhan City	China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention; BGI PathoGenesis Pharmaceutical Technology Co., Ltd
BetaCoV/Wuhan/WH19005/2019	NMDC60013002-10	NMDC		2019-12-30	China / Hubei Province / Wuhan City	China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention; BGI PathoGenesis Pharmaceutical Technology Co., Ltd
BetaCoV/Wuhan/WH19008/2019	NMDC60013002-06	NMDC		2019-12-30	China / Hubei Province / Wuhan City	China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention; BGI PathoGenesis Pharmaceutical Technology Co., Ltd
BetaCoV/Wuhan/YS8011/2020	NMDC60013002-07	NMDC		2020-01-07	China / Hubei Province / Wuhan City	China CDC; Shandong First Medical University & Shandong Academy of Medical Sciences; Hubei Provincial Center for Disease Control and Prevention; BGI PathoGenesis Pharmaceutical Technology Co., Ltd
TG13	GWHABKP0000000	Genome Warehouse	EPI_ISL_402131, MN	2013-07-24	China / Yunnan / Pu'er	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
WIV02	GWHABKK0000000	Genome Warehouse	EPI_ISL_402127, MN	2019-12-30	China/ Hubei / Wuhan	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
WIV04	GWHABKL0000000	Genome Warehouse	EPI_ISL_402124, MN	2019-12-30	China / Hubei / Wuhan	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
WIV05	GWHABKM0000000	Genome Warehouse	EPI_ISL_402128, MN	2019-12-30	China / Hubei / Wuhan	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
WIV06	GWHABKN0000000	Genome Warehouse	EPI_ISL_402129, MN	2019-12-30	China / Hubei / Wuhan	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
WIV07	GWHABKO0000000	Genome Warehouse	EPI_ISL_402130, MN	2019-12-30	China / Hubei / Wuhan	CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences
Wuhan-Hu-1	MN908947	GenBank	NC_045512	2019-12	China / Hubei Province / Wuhan City	Shanghai Public Health Clinical Center & School of Public Health, Fudan University, Shanghai, China

*The following 4 sequences were obtained from GenBank for analysis: AY572038.1_SARS_civet020,/1-29683, AY714217_SARS_coronavirus_CDC#200301157,/1-29727, MG772933_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZC45,/1-29802, MG772934_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21,/1-29732

Whole genome sequence alignment of 2019-nCoV together with human SARS, civet SARS, several bat SARS like CoV related virus supported previous conclusion that 2019-nCoV most closely clusters with a novel bat RaTG13 coronavirus (GenBank: MN996532)(*6*) and two bat SARS-like CoV strains (GenBank: MG772933 and MG772934, **Figure 1**)(*2*, *14*). However, there is very limited diversity among these sequenced 2019-nCoV genomes, in line with the notion that 2019-nCoV was a recent jump from its original host to human population. In addition, as there have been evidences that recombination frequently occurs within the family of betacoronavirus, rendering the accuracy of whole genome phylogeny analysis, we further analyzed the essential ORF1ab and Spike/S proteins. Results support the close relationship between bat RaTG13 strain and 2019-nCoV (**Figure 2**).

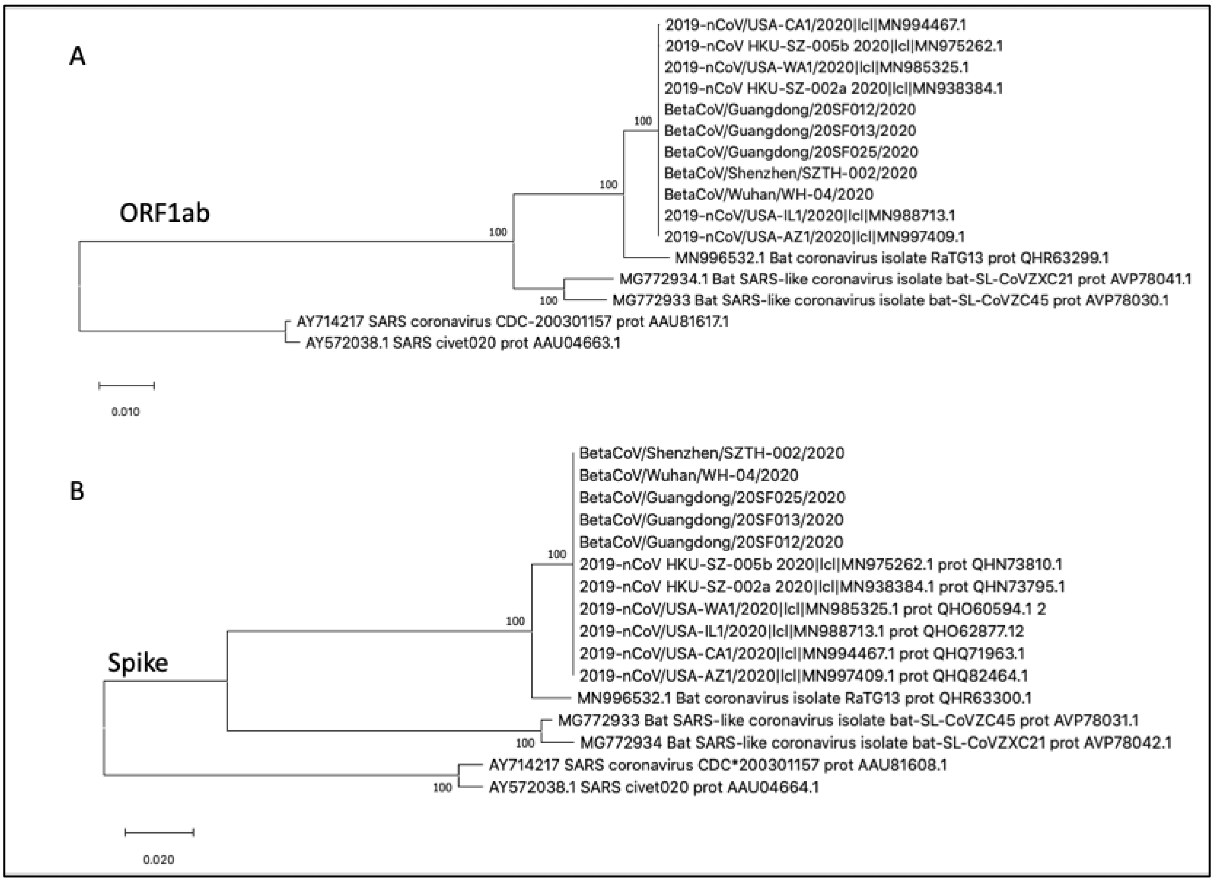
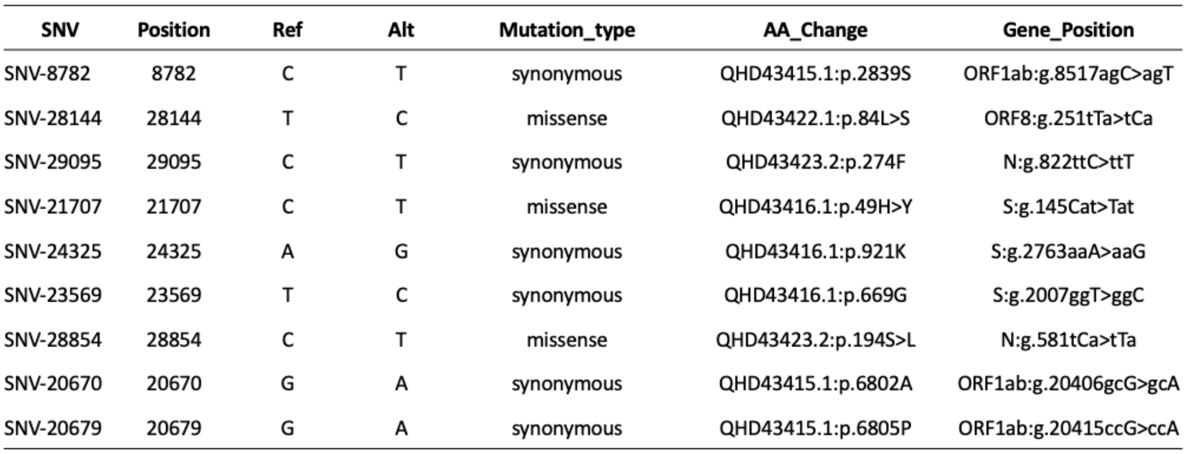


Figure 2. Phylogenetic trees reconstruction of ORF1ab and Spike/S protein. Protein sequence of eleven 2019-nCoV, RaTG13, two bat SARS-like, one civet and one human SARS viral

proteins for ORF1ab (**A**) and Spike protein (**B**) were aligned and phylogenetic tree was reconstructed with NJ with 100 bootstrap replications. SARS sequences were used to root the tree.

After careful removal of false positive mutations obtained as of 2/19/2020 from BIGD, we collected shared SNVs of each sample against the reference genome (GenBank: NC_045512). We observed that there are at least 9 SNVs, with at least two occurrences in these samples analyzed (**Table 2**).

Table 2. SNV identified with two or more occurrences



SNV	Position	Ref	Alt	Mutation_type	AA_Change	Gene_Position
SNV-8782	8782	C	T	synonymous	QHD43415.1:p.2839S	ORF1ab:g.8517agC>agT
SNV-28144	28144	T	C	missense	QHD43422.1:p.84L>S	ORF8:g.251tTa>tCa
SNV-29095	29095	C	T	synonymous	QHD43423.2:p.274F	N:g.822ttC>ttT
SNV-21707	21707	C	T	missense	QHD43416.1:p.49H>Y	S:g.145Cat>Tat
SNV-24325	24325	A	G	synonymous	QHD43416.1:p.921K	S:g.2763aaA>aaG
SNV-23569	23569	T	C	synonymous	QHD43416.1:p.669G	S:g.2007ggT>ggC
SNV-28854	28854	C	T	missense	QHD43423.2:p.194S>L	N:g.581tCa>tTa
SNV-20670	20670	G	A	synonymous	QHD43415.1:p.6802A	ORF1ab:g.20406gcG>gcA
SNV-20679	20679	G	A	synonymous	QHD43415.1:p.6805P	ORF1ab:g.20415ccG>ccA

These data were visually verified in sequence alignment (**Figure 3**). We noticed that SNV-8782, SNV-28144 and SNV-29085 were shared by these samples: BetaCoV/Guangdong/20SF012/2020|EPI_ISL_403932, BetaCoV/Guangdong/20SF013/2020|EPI_ISL_403933, BetaCoV/Guangdong/20SF025/2020|EPI_ISL_403935, MN938384.1_2019-nCoV_HKU-SZ-002a_2020, MN975262.1_2019-nCoV_HKU-SZ-005b_2020, BetaCoV/Shenzhen/SZTH-002/2020|EPI_ISL_406593, MN997409.1_2019-nCoV/USA-AZ1/2020, BetaCoV/Japan/TY-

WK-012/2020|EPI_ISL_408665, BetaCoV/Japan/TY-WK-501/2020|EPI_ISL_408666 and BetaCoV/Japan/TY-WK-521/2020|EPI_ISL_408667.

SNV-8782 and SNV-28144 were shared additionally by samples: NMDC60013002-04_BetaCoV/Wuhan/WH-04/2019, MN985325.1_2019-nCoV/USA-WA1/2020, MN994467.1_2019-nCoV/USA-CA1/2020 (**Figure 3**), and by additional strains obtained between 2/2/2020 and 2/19/2020 in BIGD database: BetaCoV/Australia/NSW01/2020|EPI_ISL_407893, BetaCoV/Australia/QLD01/2020|EPI_ISL_407894, BetaCoV/Australia/QLD02/2020|EPI_ISL_407896, BetaCoV/Belgium/GHB-03021/2020|EPI_ISL_407976, BetaCoV/Chongqing/YC01/2020|EPI_ISL_408478, BetaCoV/England/01/2020|EPI_ISL_407071, BetaCoV/England/02/2020|EPI_ISL_407073 BetaCoV/Japan/TY-WK-012/2020|EPI_ISL_408665, BetaCoV/Japan/TY-WK-501/2020|EPI_ISL_408666, BetaCoV/Japan/TY-WK-521/2020|EPI_ISL_408667, BetaCoV/Sichuan/IVDC-SC-001/2020|EPI_ISL_408484, BetaCoV/Taiwan/NTU01/2020|EPI_ISL_408489, BetaCoV/USA/IL2/2020|EPI_ISL_410045, BetaCoV/Vietnam/VR03-38142/2020|EPI_ISL_408668, BetaCoV/Yunnan/IVDC-YN-003/2020|EPI_ISL_408480.

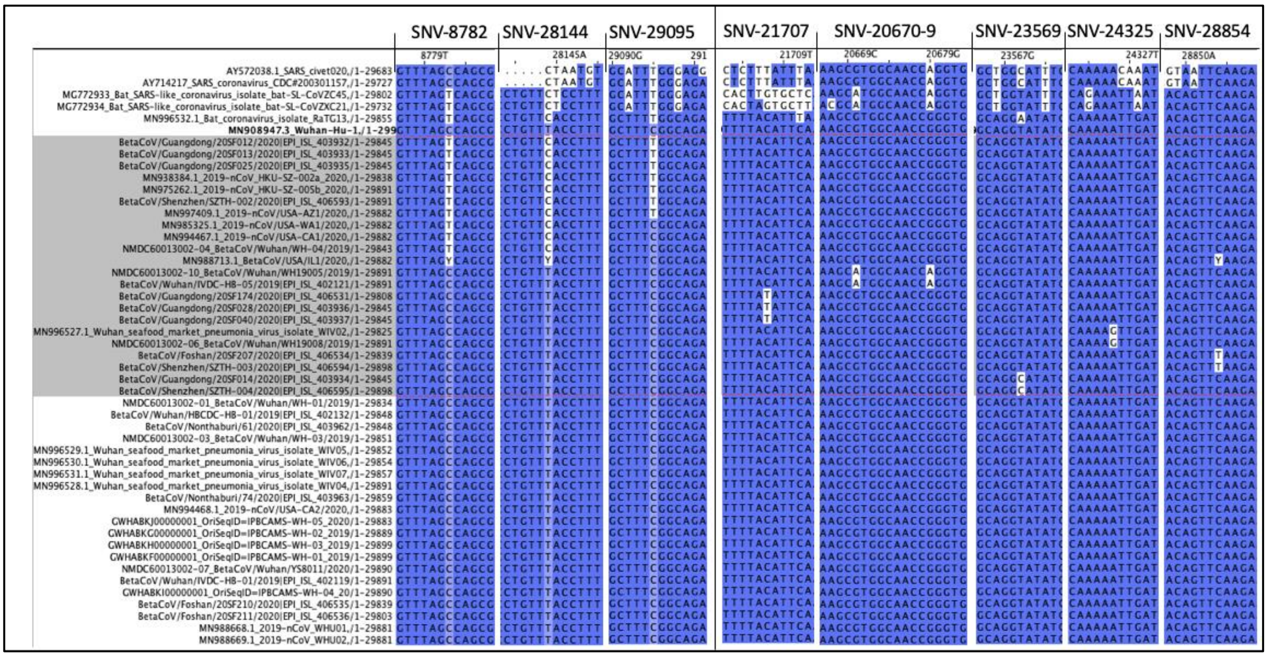


Figure 3. Sequence alignment of 2019-nCoV strains with SARS related viral genomes centered around SNV-8782, SNV-28144, SNV-29095, SNV-20670-9, SNV-21707, SNV-24325 and SNV-28854. Sequences were aligned with strain names on the left. Each SNV were labeled on the top of alignment screenshot respectively. Consensus sequence were masked blue. Number denotes the location in reference genome.

Samples of MN938384.1_2019-nCoV_HKU-SZ-002a_2020, MN975262.1_2019-nCoV_HKU-SZ-005b_2020, BetaCoV/Guangdong/20SF012/2020|EPI_ISL_403932, BetaCoV/Guangdong/20SF013/2020|EPI_ISL_403933, BetaCoV/Guangdong/20SF025/2020|EPI_ISL_403935, BetaCoV/Shenzhen/SZTH-002/2020|EPI_ISL_406593, were all collected from patients in Shenzhen, Guangdong province, indicating the infection was likely due to a local transmission in the area; Samples of NMDC60013002-04_BetaCoV/Wuhan/WH-04/2019, and internationally isolated MN985325.1_2019-nCoV/USA-WA1/2020 and MN994467.1_2019-nCoV/USA-CA1/2020 may also traced back to Wuhan travels, indicating SNV-8782 and SNV-28144 may be from parental viral ancestor and within which, SNV-29085 was derived.

The above prediction was supported that MN938384.1_2019-nCoV_HKU-SZ-002a_2020 and MN975262.1_2019-nCoV_HKU-SZ-005b_2020 that have been reported as an example of cluster transmission (*17*), indicating either fast evolution of 2019-nCoV viral genome from a single common ancestor passed down into human population, or spillover of multiple common viral ancestors from the natural reservoir of virus host to human.

We also noticed that SNV-21707 are shared by three samples collected in Zhuhai, Guangdong province by the Guangdong Provincial Center for Diseases Control and Prevention between Jan. 17 and Jan. 22 of 2020: BetaCoV/Guangdong/20SF028/2020|EPI_ISL_403936, BetaCoV/Guangdong/20SF040/2020|EPI_ISL_403937, BetaCoV/Guangdong/20SF174/2020|EPI_ISL_406531, these are likely derived from a single cluster of infected patients. Similarly, SNV-20670-9, SNV-23569, SNV-28854 and SNV-24325 were collected at the same location within a short time frame, such as SNV-24325 are shared by two samples: GWHABKK00000000-BetaCoV/Wuhan/WIV02/2019 and NMDC60013002-06_BetaCoV/Wuhan/WH19008/2019, which were collected on Dec. 30th of 2019 in Wuhan by two different labs, indicating that SNV may also have been derived from a common cluster of infected patients.

Discussion:

2019-nCoV contains a single positive single strand RNA genome of ~30kb in size. 2019-nCoV has been shown to be closely related to SARS and SARS related virus in human and in natural bat reservoir at genomic sequences and amino acid sequences of coded protein levels from phylogeny analysis.

It is expected to see that the ssRNA genome containing 2019-nCoV will evolve fast as it is transmitted among human population. Our data analysis showed, as of Feb. 2nd, we have observed multiple reoccurrence SNVs in individual positions of multiple strains along its genome. These sequence variations are unlikely to be sequencing errors and occur in multiple samples collected in different locations.

With publicly available information associated with the sequences in hand, we speculated that there may be at least two possibilities explaining this observation: (1) due to fast evolvement of viral genome in human from a single common virus ancestor, fast enough to obtain these SNVs to diverge from its common viral ancestor. (2) multiple virus ancestors were transmitted from its host to human at the time of cross species transmission. While we are unable to distinguish these two possibilities at this moment, but it will be clear with sequencing data accumulation and identification of its natural host for 2019-nCoV.

However, at this moment, the impact of these mutations among these viruses are not clear, close monitoring of the fast evolution of viral genome is needed that may aid patient care due to this pandemic coronavirus.

Conflict of interest:

The authors declare no conflicts of interest.

Acknowledgements:

The 2019-nCoV genomes were generously shared by scientists from but not limited to the following institutions: the Shanghai Public Health Clinical Center & School of Public Health, Fudan University, Shanghai, China, at the National Institute for Viral Disease Control and

Prevention, China CDC, Beijing, China, at the Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, at the Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China, at the Department of Microbiology, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China, at the Guangdong Provincial Center for Diseases Control and Prevention at the Department of Medical Sciences, at the Shenzhen Key Laboratory of Pathogen and Immunity, Shenzhen, China, at the National Institute of Health, Nonthaburi, Thailand, at the US Centers for Disease Control and Prevention, Atlanta, USA, at the Institut Pasteur, Paris, France, at the University of Melbourne, Peter Doherty Institute for Infection and Immunity, Melbourne, Australia, and at the Victorian Infectious Disease Reference Laboratory, Melbourne, Australia and additional institutions via GISAID. We gratefully acknowledge the Authors, Originating and Submitting laboratories of the genetic sequence and metadata made available through GenBank, BIGD and GISAID on which this research is based. We also thank Qian Tan and Lei Ye for their discussion on this manuscript preparation.

References:

1. Q. Li *et al.*, Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med*, (2020).
2. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*, (2020).
3. M. Bassetti, A. Vena, D. Roberto Giacobbe, The Novel Chinese Coronavirus (2019-nCoV) Infections: challenges for fighting the storm. *Eur J Clin Invest*, e13209 (2020).
4. M. L. Holshue *et al.*, First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med*, (2020).
5. C. Huang *et al.*, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*, (2020).
6. P. Zhou *et al.*, Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *bioRxiv*, 2020.2001.2022.914952 (2020).
7. L. L. Ren *et al.*, Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. *Chin Med J (Engl)*, (2020).
8. F. Wu *et al.*, A new coronavirus associated with human respiratory disease in China. *Nature*, (2020).

9. C. Xiong, L. Jiang, Y. Chen, Q. Jiang, Evolution and variation of 2019-novel coronavirus. *bioRxiv*, 2020.2001.2030.926477 (2020).
10. J. F. Chan *et al.*, Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* **9**, 221-236 (2020).
11. W. Ji, W. Wang, X. Zhao, J. Zai, X. Li, Homologous recombination within the spike glycoprotein of the newly identified coronavirus may boost cross-species transmission from snake to human. *J Med Virol*, (2020).
12. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181-192 (2019).
13. Y. Wang *et al.*, Origin and Possible Genetic Recombination of the Middle East Respiratory Syndrome Coronavirus from the First Imported Case in China: Phylogenetics and Coalescence Analysis. *mBio* **6**, e01280-01215 (2015).
14. D. Hu *et al.*, Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg Microbes Infect* **7**, 154 (2018).
15. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
16. A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp, G. J. Barton, Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191 (2009).
17. J. F. Chan *et al.*, A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*, (2020).