



## Isolation of Virus from a SARS Patient and Genome-wide Analysis of Genetic Mutations Related to Pathogenesis and Epidemiology from 47 SARS-CoV Isolates

YING ZHU, MO LIU, WEIGUANG ZHAO, JIANLIN ZHANG, XUE ZHANG, KE WANG, CHUNFANG GU, KAILANG WU, YAN LI, CONGYI ZHENG, GENGFU XIAO, HUIMIN YAN, JIAMIN ZHANG, DEYIN GUO, PO TIEN & JIANGUO WU\*

*Key Laboratory of Virology, Ministry of Education  
College of Life Sciences, Wuhan University, Wuhan 430072, China*

Received June 8, 2004; Accepted July 15, 2004

**Abstract.** Severe acute respiratory syndrome (SARS) caused by SARS-associated coronavirus (SARS-CoV) is a fatal disease. Prevention of future outbreaks is essential and requires understanding pathogenesis and evolution of the virus. We have isolated a SARS-CoV in China and analyzed 47 SARS-CoV genomes with the aims to reveal the evolution trends of the virus and provide insights into understanding pathogenesis and SARS epidemic. Specimen from a SARS patient was inoculated into cell culture. The presence of SARS-CoV was determined by RT-PCR and confirmed by electron microscopy. Virus was isolated followed by the determination of its genome sequences, which were then analyzed by comparing with other 46 SARS-CoV genomes. Genetic mutations with potential implications to pathogenesis and the epidemic were characterized. This viral genome consists of 29,728 nucleotides with overall organization in agreement with that of published isolates. A total of 348 positions were mutated on 47 viral genomes. Among them 22 had mutations in more than three genomes. Hot spots of nucleotide variations and unique trends of mutations were identified on the viral genomes. Mutation rates were different from gene to gene and were correlated well with periodical or geographic characteristics of the epidemic.

**Key words:** SARS-CoV, isolation, genomic analysis, mutation

### Introduction

In November 2002, first case of a novel infectious disease named severe acute respiratory syndrome (SARS) suddenly appeared in southern China [1]. This illness emerged and rapidly spread to different areas of Asia and then other countries around the world with a high morbidity (about 25% required intensive care) and 9.6% fatality [2]. In March 2003, the World Health Organization (WHO) made an unprecedented international effort by organizing world-leading laboratories to find the causative agent. This effort resulted in the

declaration made simultaneously by three research groups that a new SARS-associated coronavirus (SARS-CoV) was the pathogen of this disease [3–5]. When the outbreak of SARS came to an end in July 2003, it had caused a cumulative total of 8437 cases and 813 deaths worldwide [6].

Since the discovery of SARS-CoV, progresses regarding the studies of this virus have been swift dramatically as the complete viral genome was sequenced [7]. Although the definition of SARS case still largely relied on clinical and epidemiological criteria, diagnostic tests based on the detection of viral RNA and proteins have been developed [8], along with the development of vaccines [9]. Results from both phylogenetic analysis and epidemiological studies suggested the

\*Author for all correspondence:  
E-mail: wu9988@vip.sina.com

origin of SARS-CoV was animal-oriented, most likely from Himalayan palm civets, ferrets and raccoon dogs [10–13].

As a member of the Coronaviridae family, SARS-CoV is enveloped and positive-stranded RNA virus. It harbors 23 coding sequences, including 4 primary structural proteins (nucleocapsid protein N, spike protein S, membrane protein M, and small envelope protein E); 5 non-structural proteins (X1, X2, X3, X4, X5); and 1 polyprotein that compose two ORFs (ORF1a and ORF1b). Polyprotein catalytically auto-processes to produce a group of proteins including proteases (PLPpro and 3Clpro), RNA-dependent polymerase (POL), RNA helicase (HEL), and function unknown proteins [4,5,7]. Like other RNA viruses, whose most striking characteristic is the high rate of genetic mutation [11,14–18].

Despite the fact that the SARS-CoV can cause an atypical and fatal form of pneumonia, the genome structure, gene expression pattern, and protein profiles of the virus are similar to those of other conventional coronaviruses [17], which are only responsible for mild respiratory tract infections in a wide range of animals including humans, pigs, cows, mice, cats, and birds [10,19]. It is possible that distinct patterns of several genes and unique variations in the SARS-CoV genome may contribute to its severe virulence or pathogenesis. The mechanism of SARS-CoV pathogenesis may involve both direct viral cytosidal effects on the target cells and immune-mediated mechanisms. Potential mutability of the viral genome may pose problems in the control of future SARS epidemics.

In this report, we described the isolation of a new SARS-CoV strain (WHU) from a patient in Hubei Province, China during the late period of SARS outbreak. Complete genome sequence of WHU isolate was determined and compared with that of 46 other SARS-CoV strains whose complete genomic sequences were available at the time analyzed. Comparative study of genetic characterization and nucleotide variation of all known SARS-CoV offers insights into understanding functions of the viral genes and revealing the evolution trends of the virus. It would also provide basis for clinical diagnosis, future developing potential drugs and vaccines against SARS-CoV infections.

## Materials and Methods

### SARS Patient

The SARS patient was an 18-year-old male from Jiayu County, Hubei Province, China. He worked in Beijing during that time when SARS outbreak was occurring. He came back to Hubei Province and became ill on April 29th, 2003 with fever and atypical pneumonia, and was admitted to hospital for isolation and treatment on May 3rd 2003.

### Virus Isolation and cDNA Synthesis

VeroE6 cells were inoculated with specimen obtained from the SARS patient. The presence of the SARS-associated coronavirus in infected cell cultures was determined by the appearance of cytopathic effects (CPE) as well as by RT-PCR amplification using primers (Primer-1/Primer-2 and Primer-3/Primer-4; Table 1) specific to the SARS-CoV. Viral particles were examined under electron microscope. Viral RNA was extracted from infected VeroE6 cells based on the procedures described by the manufacture (Invitrogen, Carlsbad, CA).

Table 1. Primers used in this study

Primers	Sequences (from 5' to 3')
Primer 1	ATGAATTACCAAGTCAATGGTTAC
Primer 2	CATAACCAGTCGGTACAGCTAC
Primer 3	TACACACCTCAGCGTTG
Primer 4	CACGAACGTGACGAAT
Primer 5	ATATTAGGTTTTTACCTACCCAG- GAAAAGCCAACC
Primer 6	CTCTGCATCGTCTTCTTCTCCTC
Primer 7	GAGGAAGAAGAGGACGATGCAGAG
Primer 8	GAATGTGACAGATAGCTCTCGTG
Primer 9	CACGAGAGCTATCTGTACATTC
Primer 10	CATCCATAAGCACATAACGAGTGTC
Primer 11	GACACTCGTTATGTGCTTATGGATG
Primer 12	GCAGTGGCATAAGCGGCATATGATG
Primer 13	CATCATATGCCGTTATGCCACTGC
Primer 14	CATAGTACTACAGATAGAGACACCAGC
Primer 15	GCTGGTGTCTCTATCTGTAGTACTATG
Primer 16	CAACGCTGAGGTGTGTAGGTGC
Primer 17	GCACCTACACACCTCAGCGTTG
Primer 18	GTCAGCATTCCAAGAATGCTCTG
Primer 19	CAGAGCATTCTTGAATGCTGAC
Primer 20	CACTAACTAGAGCAGCAGTGTAGGC
Primer 21	GCCTACACTGCTGCTCTAGTTAGTG
Primer 22	CTGTTGTCACCTACTGTACTAGC
Primer 23	GCTAGTACAGTAAAGTGACAACAG
Primer 24	TTTGTCACTTCTCCTAAGAAGCTAT

The first strand of the viral cDNA was synthesized from extracted viral RNA by reverse transcription PCR using random primers provided by the manufacture (Promega, Madison, WI). Double-stranded DNA fragments were produced by PCR amplification of the viral cDNA using 10 pairs of specific primers (primer 5 to primer 24; Table 1) designed to cover entire viral genome based on the sequences of SARS-CoV strain HKU-39849 (accession number AY278491). Each of the PCR products was cloned into vector pGEM-T, respectively. Random clones were selected for DNA sequencing analysis. Sequences representing the entire viral genome was fully assembled and edited by DNAsis software programs. Nucleotide sequences of complete genome of the SARS-CoV isolate (WHU) were deposited to GenBank (accession number AY394850).

#### Genomic Sequence Analysis

The complete genome sequences of all 47 SARS-associated coronaviruses were downloaded from GenBank (Table 2). Homology searches for

Table 2. Accession numbers of genomic sequences of 47 SARS-associated coronaviruses released in the GenBank

Genome	Accession number	Genome number	Accession number
Fra	AY310120	SARS	NC004718
GD69	AY313906	SoD	AY461660
Sino3-11	AY485278	Sino1-11	AY485277
CUHK-AG03	AY345988	CUHK-AG02	AY345987
CHUK-AG01	AY345986	CHUK-Su10	AY282752
PUMC-03	AY357076	PUMC-02	AY357075
PUMC-01	AY350750	GZ50	AY304495
AS	AY427439	HSR-1	AY323977
Sin2774	AY283798	WHU	AY394850
HKU-39849	AY278491	GD01	AY278489
TWC3	AY362699	TWC2	AY362698
Sin2748	AY283797	Sin2679	AY283796
Sin2677	AY283795	Sin2500	AY283794
Urbani	AY278741	TWY	AP008581
TWS	AP006560	TWK	AP006559
TWJ	AP006558	TWH	AP006557
CUHK-W1	AY278554	Taiwan TC3	AY348314
Taiwan TC2	AY338175	Taiwan TC1	AY338174
TWC	AY321118	Frankfurt	AY291315
BJ04	AY279354	BJ03	AY278490
BJ02	AY278487	ZJ01	AY297028
TOR2	AY274119	TW1	AY291451
BJ01	AY278488	Shangai QXC1	AY463059
Shangai QXC2	AY463060		

the DNA sequences were conducted and their deduced amino acid sequences were analyzed through the public database with the BLAST search program provided by the National Center for Biotechnology Information (NCBI). Sequence alignment was performed using software ClustalW and further analyzed using software BioEdit.

Nucleotide sequences of the entire genome of newly identified WHU strain along with that of other 46 SARS-CoV isolates released in the GenBank were aligned with the ClustalW software program. Phylogenetic trees were created for all nucleotide sequences by neighbor-joining and parsimony methods. Sequences were analyzed with reference to the trees to reveal character states relevant to phylogenetic branching.

#### Results

During late period of the SARS outbreak in 2003, three patients were identified as probable SARS cases in Hubei Province, a less SARS representative area in China. In order to study the SARS-CoV caused disease, we obtained specimen from one of the patients. Seven days after inoculation of VeroE6 cells with patient specimens, CPE was appeared on the infected cells (Fig. 1) indicating the presence of an infectious agent. Two specific amplicons were detected by RT-PCR amplifications using extracted viral RNA as templates when two pairs of SARS-CoV specific primers were used, respectively (data not shown). These results implicated that exist of a SARS-CoV in the specimen was highly possible. Coronavirus-

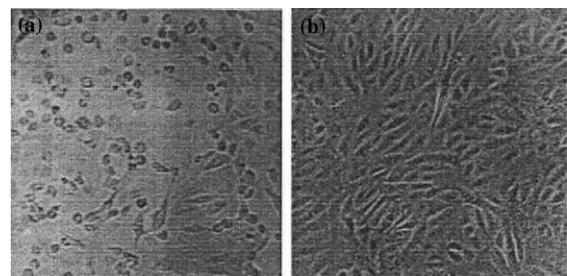


Fig. 1. Cytopathic effects (CPE) of SARS-CoV WHU strain on VeroE6 cells. VeroE6 cells were inoculated with blood samples obtained from the SARS patient (Fig. 1A) or with PBS (Fig. 1B). Seven days after inoculation, cytopathic effects (CPE) of VeroE6 cells were observed under microscope.

like particles were observed when we further examined infected cells under electron microscope (data not shown). In addition, SARS-CoV antibodies were detected from the patient's serum. All together, these results provided substantial evidence to suggest that this patient was infected by SARS-CoV, named WHU strain.

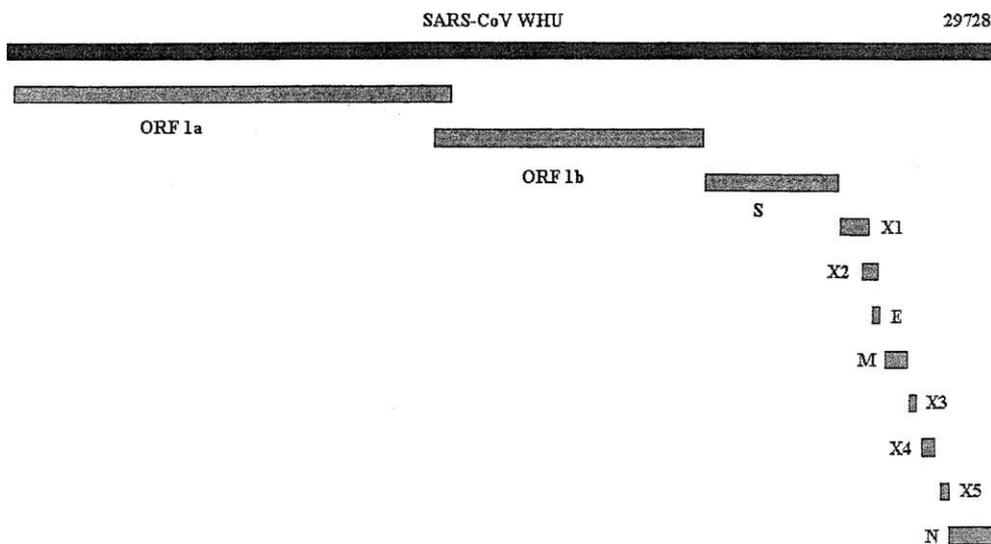
After identification of the WHU strain, we isolated the virus and determined complete nucleotide sequences of its genome (accession number AY394850). Since this virus was the only SARS-CoV that has ever been isolated and sequenced from Hubei Province, we carried out detailed sequence analysis of its entire genome. Results from sequence analysis indicated that the genome of WHU strain consisted of 29,728 nucleotides with a two-nucleotide deletion at residuals 27,825 and 27,826.

Phylogenetic analysis was conducted with the genome sequences of the WHU strain and that of all 46 SARS-CoV isolates, whose genomic sequence information was fully available in the public databases (Table 2). Both phylogenetic study and sequence analysis indicated that the overall genome organization and predicted proteins of WHU isolate were in agreement with published studies on other SARS-CoV isolates (Fig. 2). Like all SARS-

CoV isolates, the WHU strain belongs to a new group of coronavirus [3]. However, the WHU isolate with a two-nucleotide deletion was genetically diverse from most of the published SARS-CoV isolates, but closely related to TWC strain (Fig. 3).

To investigate the variations of nucleotide sequences among SARS coronaviruses, we performed a genome-wide analysis of genetic mutations on all 47 SARS-CoV genomes. Results indicated that a total of 348 positions on the 47 viral genomes had alternative nucleotides. Among them, 22 positions with mutations occurred on more than three viral genomes (Table 3, Fig. 4). The most variable spot was residual 26,477, at which nucleotides of 20 viral genomes were mutated. The second highest variable positions were residual 3852 and 11,493, since 14 viral genomes showed nucleotide changes at each of these two locations. Residual 17,564 had mutations on 9 viral genomes, while residual 9404, 19,084, 19,838, 22,222, 27,813, and 27,827 showed moderate mutability among 7 viruses. The rest 12 sites had mutations occurred on genomes ranging from 3 to 5 viruses (Table 3 and Fig. 4).

Our next step was to determine whether the high mutability had any implications linked to the viral genes or their functions (Fig. 4 and Table 3). After



*Fig. 2.* Schematic Organization of major genes on the genome of the SARS-associated coronavirus WHU strain isolated from a SARS patient from Wuhan, China. Location of individual genes on the viral genome: ORF1a (265–13,413); ORF1b (13398–21485); S gene (21,492–25,259); X gene (25268–26092); X2 gene (25,689–26,153); E gene (26,117–26,347); M gene (26,398–27063); X3 gene (27,074–27,265); X4 gene (27,273–27,641); X5 gene (27,862–28,116); N gene (28,118–29,386).

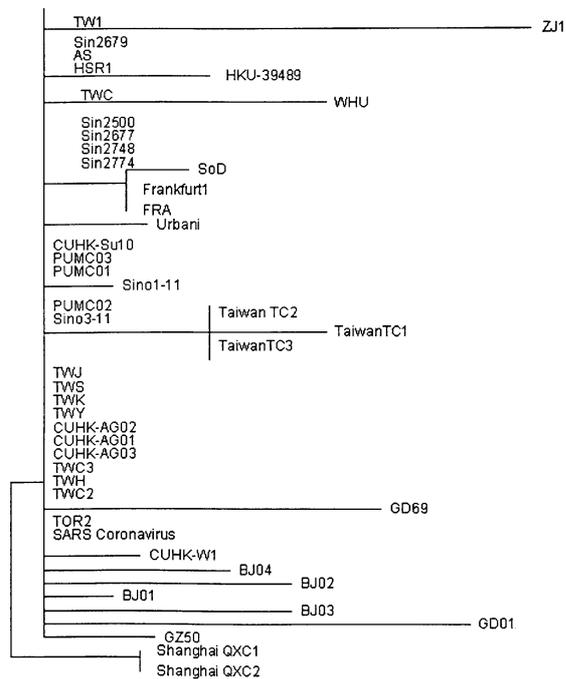


Fig. 3. Phylogenetic trees of the full-length nucleotide sequences of the genome of SARS coronavirus WHU isolate and that of other 46 SARS-CoV isolates. All sequences were aligned with the CLUSTAL-W software program. Phylogenetic trees were created for all nucleotide and predicted amino acid sequences by neighbor-joining program and parsimony methods.

further comparison and analysis of the viral sequences, we realized that polyprotein gene (ORF1 a and ORF1 b) had the highest variation rate among all genes. This region not only carried 11 mutations, but also had the second highest variable positions (residual 3852 and 11,493). ORF1b gene contains additional two residuals (17,564 and 19,084) at which 7 viruses were mutated. We also noticed that the S gene had a high mutability with residual 22222 mutated in 7 viruses, residual 21721 in 6, and residual 24933 in 3. Two positions with high mutation rate were identified within the M gene. One was located at the most variable residual 26477, at which 20 viruses were mutated. The other one was residual 26600, at which 6 viral genomes were changed. E gene and N gene had one mutation spot at residual 26203 and 28276, respectively. Among five non-structural genes, X4 had one mutation site at residual 27243 with mutation rate of 5, while X5 gene had two mutation spots at residual 27813 and 27827 with mutation rate of 7 (Fig. 4 and Table 3).

Based on the recommendations from WHO [6], all SARS cases can be divided periodically into early-period case, mid-period case, and late-period case (Table 4). In this study, we proposed all 47 known viral isolates into two groups, early-mid period and mid-late period group (Table 5). Based on results from sequence analysis, we realized that there were some correlations between genetic mutations of the virus and periodical or geographic characteristics of the outbreak. Several residuals (9404, 9854, 17564, 19838, 21721, 22222, 27243 and 27827) with mutations only occurred during early-mid period, while nucleotides at some positions (2557, 11448, 18965, 19084, 24933, 27813 and 28276) changed exclusively during mid-late period of the epidemic (Tables 3 5). In addition, some genetic mutations were linked to certain geographic regions where the viruses isolated. For instance, high genetic mutation rate at position 3852 was mainly found in viruses isolated from Taiwan. Mutations at residual 26203 occurred in most Taiwan isolates (60%), but not found in any isolates identified from other regions around the world. Moreover, all three viral strains (FRA, SoD and Frankfurt) isolated from Europe had mutations at the same residuals, 2557, 11448 and 24933, while the rest isolates showed no changes in these positions (Tables 3 and 5).

## Discussion

Although the SARS epidemic ended after 6 months spreading, many important questions remain unclear. What is the natural reservoir of SARS-CoV; where and how the virus crossed the barriers between its reservoir and human to initiate reservoir-human transmission, and subsequent human-to-human infection. It was proposed that the natural reservoir of SARS-CoV was animal originated [10,11,13], most likely Himalayan palm civets [12]. This was not a surprise, since many fatal human viruses including HIV and influenza virus were originated by transmission from animals. HIV pandemic had happened as a consequence of the combination of transmission of SIVcpz from chimpanzee and common practice of "hunting and field-dressing chimpanzee" in West Central Africa [20]. Similarly in Southern China, where SARS-CoV initially emerged, people used

Table 3. Summary of genetic mutations within genes of 47 SARS-associated coronaviruses

ORF 1a	Position	2557	3852	9404	9854	11448	11493
	Mutation rate	3	14	7	6	3	14
	Nucleotide change	G-A	T-C	T-C	C-T	C-T	C-T
	Amino acid change	A-T	F-S	V-A	A-V	Silent	Silent
ORF 1b	Position	17,564	18,965	19,064	19,084	19,838	
	Mutation rate	9	4	5	7	7	
	Nucleotide change	T-G	T-A	A-G	C-T	A-G	
	Amino acid change	N-E	Silent	Silent	T-I	Silent	
S gene	Position	21,721	22,222	23,220	24,933		
	Mutation rate	6	7	4	3		
	Nucleotide change	G-A	T-C	T-G	C-T		
	Amino acid change	G-D	I-T	S-A	L-F		
E gene	Position	26,203					
	Mutation rate	6					
	Nucleotide change	C-T					
	Amino acid change	Silent					
M gene	Position	26,477	26,600				
	Mutation rate	20	6				
	Nucleotide change	T-G	C-T				
	Amino acid change	F-C	A-V				
X4 gene	Position	27,243					
	Mutation rate	5					
	Nucleotide change	C-T					
	Amino acid change	T-L					
X5 gene	Position	27,813	27,827				
	Mutation rate	7	7				
	Nucleotide change	C-T	T-C				
	Amino acid change	S-L	C-R				
N gene	Position	28,276					
	Mutation rate	4					
	Nucleotide change	C-T					
	Amino acid change	R-W					

Positions of nucleotides were based on that of the SARS-CoV Urbani strain.

to consume wild animal meat and some of the animals are now confirmed to carry SARS-like coronavirus [12].

Another question is whether SARS outbreak will come back. At the beginning of 2004, three SARS cases were reported indicating SARS do come back. However, the situation of this year seems quite different from last year, since transmission, infection and severity of SARS-CoV were clearly weakened. One possible explanation is that it might be just a preface of SARS epidemics. Like last year, in the early period of SARS pandemics, the virus did not show strong toxicity. Another possibility is that SARS-CoV might be truly weakened due to many reasons including genetic

mutations, like the influenza FluA virus which has caused a disaster outbreak in 1918 and was weakened after the pandemic that took 20 million lives [21]. Influenza epidemics throughout the world occurred periodically between the first pandemic and present time due to the viral antigenic drift and shift. These processes also resulted in the appearance of influenza B and C virus with significant differences in genetic characterizations [22]. It would be important to find out if SARS-CoV has similar epidemic rules as influenza virus dose, whether SARS-CoV is weakening or will SARS breakout periodically. While these questions remain to be addressed, it is for sure that the SARS-CoV certainly has a high mutation rate on

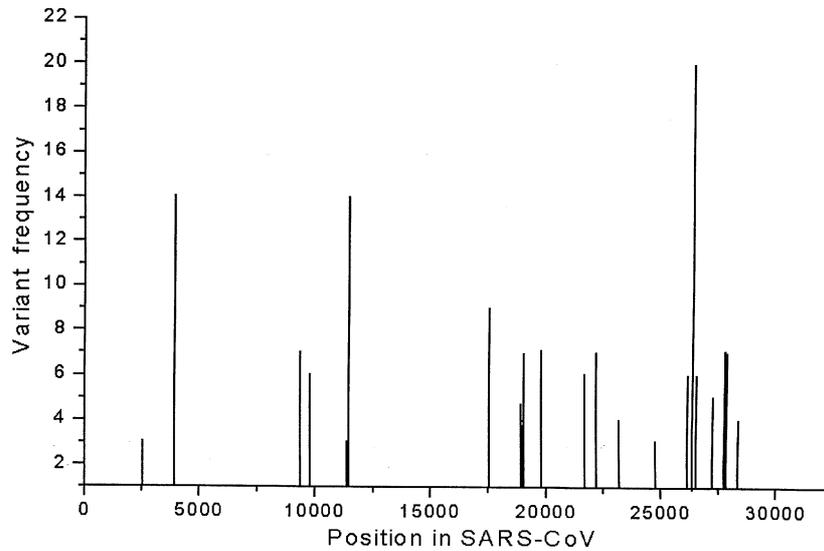


Fig. 4. Positions and mutation rates of 22 highly variable residuals on the viral genomes based on the comparison of genomic sequences of 47 SARS-CoV isolates published in the databases (GenBank).

Table 4. Classification of geographic locations and periodic divisions of the SARS outbreak from late 2002 and early 2003

Early period case	Mid-period case	Late-period case
Southern China (16 November, 2002)	Mainland China (13 February 2003)  China Hong Kong SAR (15 February 2003) Canada (23 February 2003) Vietnam (23 February 2003) Singapore (25 February 2003)	Mainland China-Taiwan (25 February 2003)  Germany (9 March 2003) Italy (12 March 2003) Russian Federation (5 May 2003)

Table 5. Division of 47 SARS-associated coronaviruses based on periods of the SARS outbreak

Early mid-period				Mid-late period					
GD 69	GZ 50	GD 01	Sino 3-11	Sin 2748	SoD	AS	HSR-1	WHU	Sin 2774
PUM C03	PUM C02	PUM C03	BJ 04	Sin 2679	TW/C3	TW/C2	TW/C	Sin 2500	Sin 2677
BJ02	BJ01	HKU-39849	CUHK-AG03	Urbani	TW/Y	TWS	TWK	TWJ	TW H
Sino 1-11	CUHK-AG01	CUHK-W1	CUHK-AG02	Taiwan TC3		Taiwan TC2		Taiwan C1	
BJ03				Frankfurt Shangai XC2		Shangai QXC1		TW-1	TOR/2
Mutations occurred at residuals 9404, 9854, 17564, 19838, 21721, 22222, 27243 and 27827				Mutations occurred at residuals 2557, 11448, 18695, 19084, 24933, 27813 and 28276					

its genome, which could in turn play significant roles in its pathogenicity and epidemics of the disease.

Molecular epidemiology and genome-wide analysis of mutations among SARS-CoV have provided insights into our understanding some of the questions [11,14–18]. For instance, except the geographic distribution of potential animal reservoirs, the high homologies among SARS-CoV of human and SARS-like coronavirus of animals strongly supported the hypothesis of animal origin of SARS-CoV [12]. It is possible that some mutations on the viral genome were responsible for the transmission of SARS-CoV from animals to human.

In an effort to study the SARS-CoV, we identified and genetically sequenced a new SARS-CoV isolated from a patient with SARS in Hubei Province. Hubei was a less SARS representative area in China, because there were only a total of three patients confirmed as probable SARS cases and only one viral strain was isolated from this region. These facts prompted us to study this virus further. Our sequence analysis indicated that although the overall genome organization of WHU (Fig. 2) is in agreement with published studies on other isolates, WHU carried a two-nucleotide deletion at residuals 27825 and 27826 was genetically diverse from most SARS-CoV isolates. These results implicated that mutations occurred during the viral transmission from Beijing to Hubei, although we do not know at this point whether these mutations have any biological significance. It is interesting to notice that although the SARS-CoV virus evaded human population only for 6 months, its genetic information already altered in many ways during its short journey of human transmission.

Individual viral genes displayed distinct patterns of genetic mutations at different time during the SARS outbreak. For instance, mutability of the S gene was high during early-mid period, but low during mid-late period of the epidemic, which suggested that mutability of S gene decreased as viral transmission increased. One possible explanation for this observation is that during early-mid period of the epidemic, as the gene encoding protein for the recognition of receptors of the host and for the mediation of viral entry into host cells, S gene had to change at a high frequency in order

to quickly fulfill its biological roles. Once the viral adaptation to human cells completed or reached its equilibrium, genetic changes were less important or no longer needed. Thus, genetic information of S gene became relatively stable during mid-late period of the outbreak [23]. Another example is ORF lab that encodes the polyprotein of SARS-CoV. Like S gene, ORF lab was also actively involved in genetic mutations. However, in contrast to S gene, mutability of ORF lab was low at the beginning, but high during mid-late period of the epidemics. This observation can be explained well by the fact that the toxicity of SARS-CoV was weakened in mid-late period. Other structural genes including E, M, and N genes were more conserved at beginning of the outbreak, but underwent genetic changes at the end of transmission. This pattern of genetic mutation obviously reflects biological roles of these structural genes in viral particles assembly, which in turn crucial for the virus to fight with increasing immune pressures from the hosts. Genetic analysis of non-structural genes showed that they intended to keep genetic information conserved throughout the entire process of transmission. Therefore, these genes may prove to be ideal targets for the diagnosis of SARS-CoV, screening antiviral drugs, and perhaps developing antiviral vaccines.

Patterns of genetic mutations of certain viral genes were linked to geographic locations from where the virus isolated. Mutations at residuals 3825 and 26203 within the X5 and E genes could clearly set the Taiwan isolates apart from others. Thus, these two positions may be used as molecular signatures in the identification of Taiwan isolates. Similar phenomena were also found in three viral strains (SoD, FRA, and Frankfurt) isolated from Europe during mid-late period of the outbreak. These viral strains had mutations at the same residuals (2557, 11,448 and 24,933), while all isolates from other regions did not show any changes at these positions. This kind of specific mutation pattern may reflect relatively independent geographical locations of Taiwan and Europe. We speculated that population in these regions perhaps developed unique immunity due to their unique locations, for which the virus had to make specific genetic mutations in order to invade these populations. In addition, based on genome-wide

mutation analysis, some viral strains isolated from Beijing had a close relationship to isolates identified from Southern China during early-mid period of the outbreak. It could be translated to that at least these SARS-CoV isolates found in Beijing were originally from Southern China.

Much have to be done in order to understand thoroughly the evolution, transmission, origin, and infection of SARS-associated coronavirus. It is interesting to recognize that genome-wide mutation analysis could provide new insights into our understanding the route of viral transmission and predication or perhaps prevention of future SARS epidemics. Our study would provide a rational and hypothesis-driven approach to study these questions, develop rapid diagnostic tests, and design measurement to prevent this fatal disease. In addition, fully understand molecular mechanism of genetic mutations would provide insights into understanding plausible transmission route of SARS-CoV from animal to humans as well as from human to human, and trends of changing in pathogenicity of SARS-CoV during its rout of transmission and path of evolution.

### Acknowledgment

This research was supported by the SARS Special Grant of Wuhan University.

### References

- Peng G.W., He J.F., Lin J.Y., Zhou D.H., Yu D.W., Liang W.J., Li L.H., Guo R.N., Luo H.M., and Xu R.H., *Chin J Epidemiol* 24, 350–352, 2003.
- WHO. Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS). [http://www.who.int/csr/sarscountry/2003\\_07\\_11/en](http://www.who.int/csr/sarscountry/2003_07_11/en).
- Ksiazek T.G., Erdman D., Goldsmith C.S., Zaki S.R., Peret T., Emery S., Tong S., Urbani C., Comer J.A., Lim W., Rollin P.E., Dowell S.F., Ling A.E., Humphrey C.D., Shieh W.J., Guarner J., Paddock C.D., Rota P., Fields B., DeRisi J., Yang J.Y., Cox N., Hughes J.M., LeDuc J.W., Bellini W.J., and Anderson L.J., SARS Working Group, *N Engl J Med* 348, 1953–1966, 2003.
- Drosten C., Gunther S., Preiser W., van der Werf S., Brodt H.R., Becker S., Rabenau H., Panning M., Kolesnikova L., Fouchier R.A., Berger A., Burguiere A.M., Cinatl J., Eickmann M., Escirou N., Grywna K., Kramme S., Manuguerra, J.C., Muller S., Rickerts V., Sturmer M., Vieth S., Klenk H.D., Osterhaus A.D., Schmitz H., and Doerr H.W., *N Engl J Med* 348, 1967–1976, 2003.
- Rota P.A., Oberste M.S., Monroe S.S., Nix W.A., Campagnoli R., Icenogle J.P., Periaranda S., Bankamp, B., Maher K., Chen M.H., Tong S., Tamin A., Lowe L., Frace M., DeRisi J.L., Chen Q., Wang D., Erdman D.D., Peret T.C., Burns C., Ksiazek T.G., Rollin P.E., Sanchez A., Liffick S., Holloway B., Limor J., McCaustland K., Olsen-Rasmussen M., Fouchier R., Gunther S., Osterhaus A.D., Drosten C., Pallansch M.A., Anderson L.J., and Bellini W.J., *Science* 300, 1394–1399, 2003.
- Department of Communicable Disease Surveillance and Response. WHO Consensus document on the epidemiology of severe acute respiratory syndrome (SARS). <http://www.who.cds/csr/gar/2003.11>.
- Marra M.A., Jones S.J., Astell C.R., Holt R.A., Brooks-Wilson A., Butterfield Y.S., Khattra J., Asano, J.K., Barber S.A., Chan S.Y., Cloutier A., Coughlin S.M., Freeman D., Gim N., Griffith O.L., Leach S.R., Mayo M., McDonald H., Montgomery S.B., Pandoh P.K., Petrescu A.S., Robertson A.G., Schein J.E., Siddiqui A., Smailus D.E., Stott J.M., Yang G.S., Plummer F., Andonov A., Artsob H., Bastien N., Bernard K., Booth T.F., Bowness D., Czub M., Drebot M., Fernando L., Flick R., Garbutt M., Gray M., Grolla A., Jones S., Feldmann H., Meyers A., Kabani A., Li Y., Normand S., Stroher U., Tipples G.A., Tyler S., Vogrig R., Ward D., Watson B., Brunharm R.C., Krajden M., Petric M., Skowronski D.M., Upton C., and Roper R.L., *Science* 300, 1399–1404, 2003.
- Poon L.L., Chan K.H., Wong O.K., Yam W.C., Yuen K.Y., Guan Y., Lo Y.M., and Peiris J.S., *J Clin Virol* 28(3), 233–238, 2003.
- Gao W., Tamin A., Soloff A., D'Aiuto L., Nwanegbo E., Robbins P.D., Bellini W.J., Barratt-Boyes S., and Gambotto A., *Lancet* 362(9399), 1895–1856, 2003.
- Enserink M., *Science* 300, 1351, 2003.
- Eickmann M., Becker S., Klenk H.D., Doerr H.W., Stadler K., Censini S., Guidotti S., Massignani V., Scarselli M., Mora M., Donati C., Han J.H., Song H.C., Abrignani S., Covacci A., and Rappuoli R., *Science* 302(5650), 1504–1505, 2003.
- Guan Y., Zheng B.J., He Y.Q., Liu X.L., Zhuang Z.X., Cheung C.L., Luo S.W., Li P.H., Zhang L.J., Guan Y.J., Butt K.M., Wong K.L., Chan K.W., Lim W., Shortridge K.F., Yuen K.Y., Peiris J.S., and Poon L.L., *Science* 302(5643), 276–278, 2003.
- Ng S.K., *Lancet* 362(9383), 570–572, 2003.
- Stavrinides J., Guttman D.S., *J. Virol.* 78(1), 76–82, 2004.
- Chim S.S., Tsui S.K., Chan K.C., Au T.C., Hung E.C., Tong Y.K., Chiu R.W., Ng E.K., Chan P.Y., Chu C.M., Sung J.J., Tam J.S., Fung K.P., Waye M.M., Lee C.Y., Yuen K.Y., and Lo Y.M., *Lancet* 362(9398), 1807–1808, 2003.
- Ruan Y.J., Wei C.L., Ee A.L., Vega V.B., Thoreau H., Su S.T., Chia J.M., Ng P., Chiu K.P., Lim L., Zhang T., Peng C.K., Lin E.O., Lee N.M., Yee S.L., Ng L.F., Chee R.E., Stanton L.W., Long P.M., and Liu E.T., *Lancet* 361(9371), 1779–1785, 2003.

17. Snijder E.J., Bredenbeek P.J., Dobbe J.C., Thiel V., Ziebuhr J., Poon L.L., Guan Y., Rozanov M., Spaan W.J., and Gorbalenya A.E., *J Mol Biol* 331(5), 991–1004, 2003.
18. Tsui S.K., Chim S.S., and Lo Y.M., Chinese University of Hong Kong Molecular SARS Research Group, *N Engl J Med* 349(2), 187–188, 2003.
19. Zhong N.S., Zheng B.J., Li Y.M., Poon, Xie Z.H., Chan K.H., Li P.H., Tan S.Y., Chang Q., Xie J.P., Liu X.Q., Xu J., Li D.X., Yuen K.Y., Peiris, and Guan Y., *Lancet* 362, 1353–1358, 2003.
20. Richter M., *Can HIV AIDS Policy Law Rev* 8(1), 14–19, 2003.
21. Zambon M.C., *J. Antimicrobial Chemother* 44, Suppl B, 3–9, 1999.
22. Bush R.M., Smith C.B., Cox N.J., and Fitch W.M., *Proc Natl Acad Sci USA* 97, 6974–6980, 2000.
23. Chinese SARS Molecular Epidemiology Consortium, *Science* 303(5664), 1666–1669, 2004.