Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Short communication

Genetic distance of SARS coronavirus from the recent natural case

Zhi-Gang Wang a,*, Shu-Ping Xu b, Yan-Jun Zhang a, Qi-Yu Bao c

a Zhejiang Provincial Center for Disease Prevention and Control, 17 Laozhedazhi Road, Hangzhou 310009, China
b College of Biosystems Engineering and Food Science, University of Zhejiang, Hangzhou 310029, China
c Institute of Biomedical Informatics, Wenzhou Medical College, Wenzhou 325000, China

Received 14 June 2005; received in revised form 26 July 2006; accepted 10 October 2006

Abstract

Phylogenetic analysis of SARS coronavirus isolates based on the spike gene and protein sequence using Neighbor-Joining, maximum likelihood and Bayesian inference methods indicated that a recent human SARS-CoV isolate was closer to some human SARS-CoV isolates from earlier epidemic phase than to the SARS-CoV-like viruses isolated from wild animals during previous epidemic phase. A reasonable judgment based on phylogenetic relationship and sequence variations it is likely that the recent human SARS-CoV isolate is closer to an unknown SARS-CoV predecessor.

© 2006 Elsevier B.V. All rights reserved.

Keywords: SARS coronavirus; Phylogeny; Human and palm civet; Spike gene sequence

1. Introduction

SARS coronavirus (SARS-CoV) phylogeny and genotyping studies have processed since SARS emergence (Ruan et al., 2003; Tsui et al., 2003; Zhao et al., 2004). Genotype C and T were first suggested at the end of May 2003 (Li et al., 2003), and were further improved and named as the Yexin genotype and Xiaohong genotype (Wang et al., 2004). SARS virus was supposed to be transmitted from the wild animal to human being. This hypothesis was then supported by identification of a SARS-CoV-like virus in wild animals, such as palm civet, sold in markets in south China, for it had more than 99% of sequence identity to the SARS-CoV (Guan et al., 2003), and indicating that the virus could have recently transferred its hosts, from animals to human beings. However, recent reports indicated that SARS-CoV was distinct from the virus in palm civet and no direct evidence so far to demonstrate if the palm civet virus was the origin of the SARS-CoV or if palm civets were also infected from other species (Stadler et al., 2003). Although unlikely, the possibility that humans infected these SARS-CoV positive animals cannot be formally excluded, and there was a report of SARS-CoV transmitted from human to pig (Chen et al., 2005). Where is the SARS-CoV-like virus of palm civet...
in the chain? Are they getting it from another animal? Are palm civets infecting rodents as well as humans? These still are not known exactly.

Here, the evolitional relationship among the previous epidemic and newly occurred (WHO, 2004) (at the end of 2003) SARS-CoVs, and the previous epidemic SARS-CoV-like viruses of animal source are analyzed.

2. Materials and methods

The complete spike glycoprotein gene sequences of SARS-CoVs or SARS-CoV-like viruses download from NCBI GenBank database. GenBank accession numbers see Table 1. Multiple sequences of nucleic acid or amino acid were aligned by ClustalW 1.83 (Thompson et al., 1994). Phylogenetic trees were constructed by MEGA3.1 (Kumar et al., 2004) (Neighbor-Joining (NJ) for gene and protein sequences), PAUP* 4.0b2 (Swofford, 2002) (Maximum likelihood for gene sequences) and MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) (Bayesian inference for gene and protein sequences). The FIPV-X06170 (Feline infectious peritonitis virus) was used as an outgroup within the spike gene data set. SARS-CoV and FIPV are known to be highly identical.
Table 1
The variant locations and substitution types in the spike protein sequences

| Accession No. | Isolates  | 192 | 194 | 196 | 198 | 200 | 202 | 204 | 206 | 208 | 210 | 212 | 214 | 216 | 218 | 220 | 222 | 224 | 226 | 228 | 230 | 232 | 234 | 236 | 238 | 240 | 242 | 244 | 246 | 248 | 250 |
|---------------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| AY525636      | GD03T0013  | N   | T   | N   | S   | S   | S   | A   | T   | F   | R   | S   | S   | S   | E   | G   | T   | S   | D   | V   | R   | P   | G   | L   | E   | V   | R   | D   | D   | T   | N   | F   | R   | N   |
| AY304486      | SZ3        | K   | K   | K   | P   | L   | V   | A   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY304487      | SZ13       | K   | K   | K   | P   | L   | V   | A   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY304488      | SZ16       | K   | K   | K   | P   | L   | V   | A   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY304489      | SZ1        | K   | K   | K   | P   | L   | V   | A   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY390556      | GZ02       | T   | F   | T   | L   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   |
| AY394996      | ZS-B       | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY394981      | HGZ5811-A  | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY394988      | JMD        | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY278487      | BJ02       | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY278488      | BJ01       | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY278554      | CUHK-W1    | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY310120      | FRA        | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY274119      | TOR2       | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY283798      | Sin2774    | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY286320      | ZJ01       | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY429079      | BJ302-8    | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
| AY322207      | LY         | S   | T   | F   | T   | L   | K   | R   | R   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   | D   | A   | L   | N   | A   | Q   | L   | D   | A   |
throughout the spike gene sequence (Stavrinides and Guttman, 2004).

3. Results

Neighbor-Joining (NJ) trees of the spike gene sequences (Fig. 1) indicated that the SARS-CoV (GD03T0013) of newly occurred case is closer to the human SARS-CoVs detected in the previous epidemic early phase (such as GZ02, GZ-B and BJ02, etc) than to the palm civet or raccoon dog SARS-CoV-like viruses (SZ1, SZ3, SZ13 and SZ16) detected in the previous epidemic. The $p$-distances of new isolate (GD03T0013) with GZ02 and GZ-B isolates are smaller than with SZ3 isolate (Fig. 1). The similar results of gene sequence analysis were obtained from other models of phylogenetic trees such as Bayesian inference and Maximum likelihood methods (Fig. 2). The phylogenetic trees of the spike protein sequences were also showed similar characteristics as above (Fig. 3). The variant locations in the spike protein and the substitution types occurring in these isolates were markedly different between the human SARS-CoVs and the animal-origin SARS-CoV-like viruses comparing with the new isolate, the former has 5–7 mutual variant locations, the latter 8–9 mutual variant locations (Table 1).

4. Discussion

Our analysis is markedly different from a conclusion of previous report in SCIENCE journal (Zhao et al., 2004) (They claim that phylogenetic analysis of this S gene sequence with those from the human SARS-CoV and palm civet SARS-like coronavirus indicated that this most recent case of SARS-CoV (GD03T0013) is much closer to the palm civet SARS-like coronavirus than to any human SARS-CoV detected in the previous epidemic). Their conclusion was cited by lately reports (Song et al., 2005; Wu et al., 2004). Here, our opinion, the phylogenetic relationship should be cautiously interpreted.
In this context, as a reasonable judgment based on phylogenetic relationship and sequence variations it is likely that the recent human SARS-CoV isolate is closer to an unknown SARS-CoV predecessor than the SARS-CoVs from human or SARS-CoV-like viruses from palm civet both detected in the previous epidemic.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China (30671144).

References

Chen, W., Yan, M., Yang, L., et al., 2005. SARS-associated coronavirus transmitted from human to pig. Emerg. Infect. Dis. 11, 446–448.

Wu, C.Y., Jan, J.T., Ma, S.H., et al., 2004. Small molecules targeting severe acute respiratory syndrome human coronavirus. Proc. Natl. Acad. Sci. U.S.A. 101, 10012–10017.

Guan, Y., Zheng, B.J., He, Y.Q., et al., 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302, 276–278.

Song, H.D., Tu, C.C., Zhang, G.W., et al., 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc. Natl. Acad. Sci. U.S.A. 102, 2430–2435.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment Briefings in Bioinformatics 5, 150–163.

Li, L.J., Wang, Z.G., Lu, Y.Y., et al., 2003. Severe acute respiratory syndrome-associated coronavirus genotype and its characterization. Chin. Med. J. 116, 1288–1292.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Ruan, Y.J., Wei, C.L., Ee, A.L., et al., 2003. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet 361, 1779–1785.

Stadler, K., Massignani, V., Eickmann, M., et al., 2003. SARS—beginning to understand a new virus. Nat. Rev. Microbiol. 1, 209–218.

Stavrinides, J., Guttmann, D.S., 2004. Mosaic evolution of the severe acute respiratory syndrome coronavirus. J. Virol. 78, 76–82.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22, 4673–4680.

Tsui, S.K.W., Chim, S.S.C., Dennis, Y.M.L., et al., 2003. Coronavirus genomic-sequence variations and the epidemiology of the severe acute respiratory syndrome. N. Engl. J. Med. 349, 187–188.

Wang, Z.G., Li, L.J., Luo, Y., et al., 2004. Molecular biological analysis of genotyping and phylogeny of severe acute respiratory syndrome associated coronavirus. Chin. Med. J. 117, 42–48.

WHO. Review of probable and laboratory-confirmed SARS cases in southern China. http://www.who.int/csr/don/2004_01_27/en/.

Zhao, G.P., He, J.F., Peng, G.W., et al., 2004. Molecular evolution of the coronavirus during the course of the SARS epidemic in China. Science 303, 1666–1669.