SARS-Associated Coronavirus Quasispecies in Individual Patients

TO THE EDITOR: Quasispecies are known in RNA viruses such as hepatitis C virus and human immunodeficiency virus. Owing to poor fidelity of RNA polymerases, RNA-virus populations typically contain genetic variants that form a heterogeneous virus pool. The severe acute respiratory syndrome (SARS)—associated coronavirus, as a newly identified RNA virus, however, has been reported with relatively limited variations, and no published data have recorded the existence of quasispecies.

During the SARS outbreak from March to June in 2003, 132 patients with SARS were treated in our unit, including those with the first cluster of cases in the Beijing, China, area. We sequenced 28 full-length spike (S) glycoprotein genes of the SARS-associated coronavirus from 19 individual hospitalized patients. Viral RNA was directly extracted from clinical samples, including plasma, throat swabs, sputum, and stool. The full-length S gene was amplified as six overlapping fragments by means of a nested reverse-transcriptase polymerase chain reaction (RT-PCR). Both a TA-cloning assay and direct screening of PCR products were performed. The sequencing results were verified in three independent experiments with the use of different RT-PCR products and were confirmed by the use of platinum Pfx DNA polymerase, if necessary.

A total of 107 sequence variations with 9 recurrent variant sites were identified in analyzed sequences compared with the S gene of the BJ01 strain (GenBank accession number AY278488), including 7 nonsynonymous variants (21494 C→T, 21702 A→G, 21858 A→T, 22908 A→G, 23198 T→C, 24018 A→T, 24540 A→G [numbered on the basis of the full-length genomic sequence]). With the exception of one site (position 21702), the variant sites were first documented in humans, so far as we know.

We speculate that the higher frequency of variations in the S gene than in previous reports might be due to a broader sample collection over a longer period of time. In particular, the coexistence of sequences with and those without substitutions (with BJ01 as the reference strain) was observed in 7 of the 19 subjects. In one subject, the variant and reference sequences for three variant sites coexisted (Fig. 1A).

Phylogenetic analysis based on the S gene of the SARS-associated coronavirus showed that the new...
Identified variant sequences are closest to the isolates from the Beijing and Guangdong areas in China (Fig. 1B). Taken together, our observations suggest that the SARS-associated coronavirus may consist of complex and dynamic distributions of mutants in vivo, rather than a single, defined genomic sequence — this is a characteristic typical of RNA-virus quasispecies.

Dongping Xu, M.D.
Zheng Zhang, Ph.D.
Fu-Sheng Wang, M.D., Ph.D.
Beijing 302 Hospital
Beijing 100039, China
fswang@public.bta.net.cn