

Additional file 1 – Sequence relationship

Figure S1 shows the phylogeny reconstructed using the neighboring-joining method with mid-point rooting. One striking feature is that GZ01 was distantly separated from other sequences. All the sequences from mainland China (GZ01 and BJ01-04) clustered together and separated from the remaining sequences. Interestingly, CUHK-W1 shared some common variants with those from mainland China, suggesting multiple sources of SARS-CoV in Hong Kong. A maximum parsimony tree was similar to that in Figure S1.

The phylogeny in Figure S1 is overall consistent with the history of epidemic. The earliest occurrence of SARS was reported in the Guangdong province of China and an index patient A infected a number of people in Hong Kong on February 21, 2003 who subsequently caused outbreaks in several countries/regions [1]. The phylogeny was also supported by considering the common variants alone. Among 18 non-singletons, nine of which could be assigned to the internal branches. Five variants located at the internal branches leading to the sequences (except CUHK-W1) derived from the index patient A.

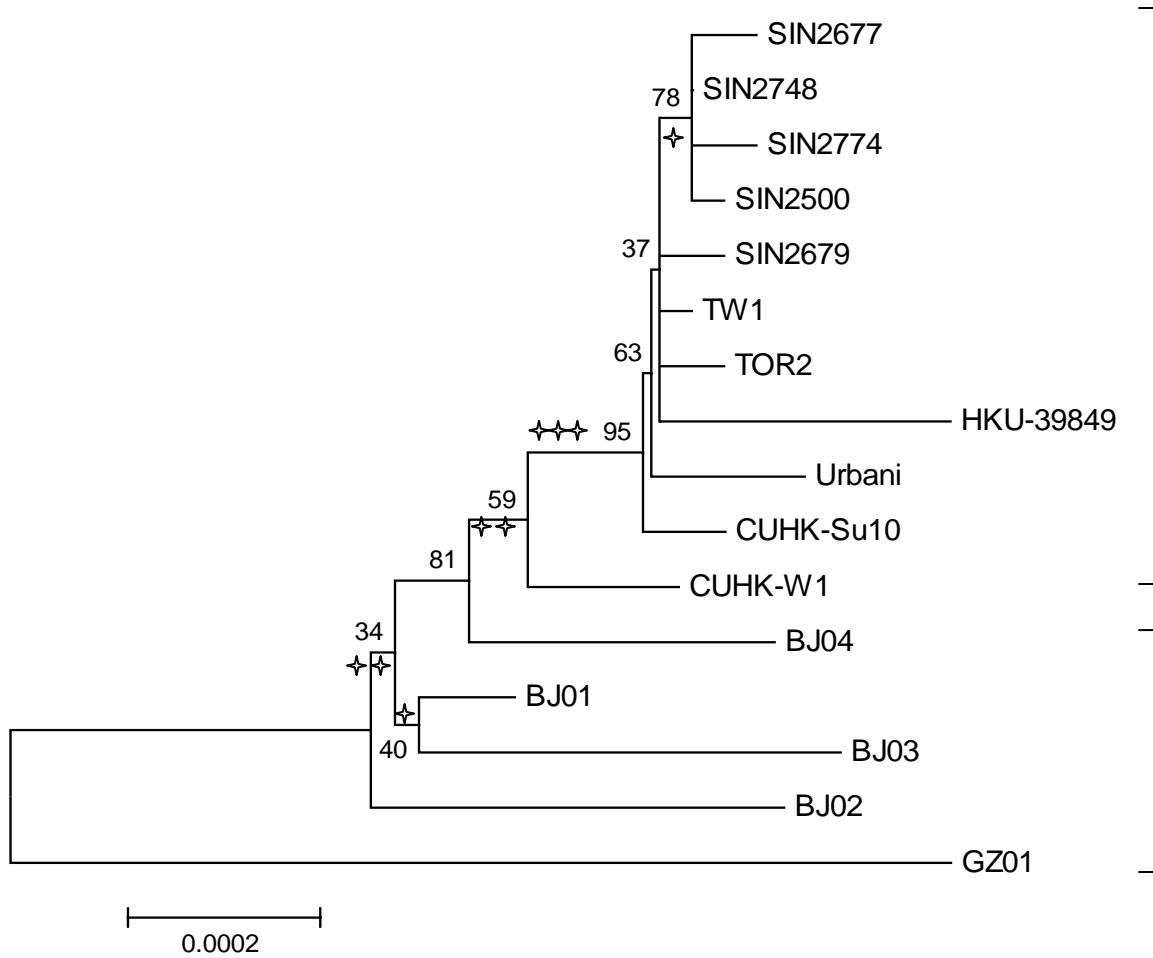


Figure S1

Phylogenetic relationships among 16 SARS-CoV genomic sequences reconstructed by neighbor-joining method. Sequence information is shown in table 4. Nine of the 18 common variants could be assigned to internal branches (labeled by a star). Five common variants (sites 17564, 19838, 22222, 27243, and 27827) located at two branches to separate the sequences isolated in mainland China (GZ01 and BJ01-04) from the others (except CUHK-W1) derived from the index patient A. Numbers on the branches indicate the bootstrap values from 1,000 replicates.

References

1. Centers for Disease Control and Prevention: **Update: Outbreak of severe acute respiratory syndrome--worldwide, 2003**. *MMWR Morb Mortal Wkly Rep* 2003, **52**:241-248.