

# SARS-CoV-2

Subjects: **Pathology**

Contributor: Shan Gao

In December of 2019, the first few cases of novel coronavirus-infected pneumonia were reported in Wuhan, China. Since then, in a series of novel reports, a research group from Nankai University of China (the Nankai group) presented several important findings of the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2): (1) The alternative translations of Nankai CDS (a 465- or 468-bp genomic region) could produce more than 17 putative proteins of the betacoronavirus subgroup B (BB coronavirus); (2) A furin protease cleavage site (FCS) was discovered in the Spike (S) protein of SARS-CoV-2; (3) 5' UTR barcoding can be used for the detection, identification, classification and phylogenetic analysis of, but not limited to coronavirus; (4) The FCS in the SARS-CoV-2 genome was acquired through the combination of copy number variations of short tandem repeats and single nucleotide substitutions [4]; and (5) two criteria were proposed to determine the intermediate host(s).

5' UTR

Barcode

Coronavirus

## 1. Introduction

In December of 2019, the first few cases of novel coronavirus-infected pneumonia were reported in Wuhan, China. Since then, in a series of novel reports, a research group from Nankai University of China (the Nankai group) presented several important findings of the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2): (1) The alternative translations of Nankai CDS (a 465- or 468-bp genomic region) could produce more than 17 putative proteins of the betacoronavirus subgroup B (BB coronavirus) [1]; (2) A furin protease cleavage site (FCS) was discovered in the Spike (S) protein of SARS-CoV-2 [2]; (3) 5' UTR barcoding can be used for the detection, identification, classification and phylogenetic analysis of, but not limited to coronavirus [3]; (4) The FCS in the SARS-CoV-2 genome was acquired through the combination of copy number variations of short tandem repeats and single nucleotide substitutions [4]; and (5) two criteria were proposed to determine the intermediate host(s) [4].

## 2. Current status

Currently, most of the researchers simply use the complete genome or gene sequences to investigate coronavirus (e.g. phylogenetic analysis) without considering the functions of the products from coronavirus genes. To overcome this shortcoming, the Nankai group proposed the joint analysis of the molecular function and phylogeny, and applied it in the study of genomes of BB coronavirus. Their analytical results showed that SARS-CoV-2 with large differences from SARS-CoV, may originate from BB coronaviruses in bats. In addition, the genotyping of 13 viruses using the 17 putative proteins revealed the high mutation rate and diversity of BB coronavirus.

The Nankai group reported for the first time (on January 21st, 2020) a very important mutation in the S proteins of betacoronavirus. By this mutation, SARS-CoV-2 acquired a furin protease cleavage site (FCS) in its S protein, which is not present in the S proteins of most other betacoronavirus (e.g. SARS-CoV). This FCS may increase the efficiency of virus infection into cells, making SARS-CoV-2 has significantly stronger transmissibility than SARS-CoV. The infection mechanism of SARS-CoV-2 may be changed to being more similar to those of MHV, HIV, Ebola virus (EBoV) and some avian influenza viruses, other than those of most other betacoronavirus (e.g. SARS-CoV). In addition, they unexpectedly found that some avian influenza viruses acquired a FCS by the similar mutation as SARS-CoV-2. Therefore, the natural mutation can result in a short insertion to form a FCS. The FCS contains the "CGGC" sequence encoding two arginine (R) residues. "CGG", however, is a rare codon for human. So these two codons were present in the SARS2-like betacoronavirus before they transmitted into human and the intermediate host(s) are mammals with a high relative frequency of "CGG" usage. Future studies of this mutation will help to reveal the stronger transmissibility of SARS-CoV-2 and lay foundations for vaccine development and drug design of, but not limited to SARS-CoV-2.

Using 5' UTR barcodes, 1,265 betacoronaviruses were clustered into four classes, and viruses in each class had similar virulence. The class 1, 2, 3 and 4 match the subgroup C, B, A and D of betacoronavirus, respectively. In particular, SARS-CoV-2 and SARS-CoV have the same 5' UTR barcode GAAAGGTAAG(ATG), which laid foundation to rename 2019-nCoV as SARS-CoV-2. In addition, the Nankai group found that Internal Ribosome Entry Sites (IRESs) may have an important role in the virulence of betacoronavirus. This important finding was reported, for the first time, to understand the virulence of SARS-CoV-2 at the molecular level. This finding can be used directly for vaccine development and design of drugs against SARS-CoV-2, but such development is not limited to coronavirus only.

Based on the mutation data of animal mitochondrion and viruses, the Nankai group proposed the non-neutral theory of molecular evolution. In this theory, there exist a substantial number of non-neutral (beneficial or deleterious) mutations in animal mitochondrion and viruses within an individual animal that are affected by natural selection. Coronavirus has such a high mutation rate that the FCS in the SARS-CoV-2 genome can be acquired in a short period. They established a mutation model, where the FCS in the SARS-CoV-2 genome was acquired through the combination of copy number variations of short tandem repeats and single nucleotide substitutions. To validate this model, they found a 12-bp insertion that resulted in formation of the FCS in a avian influenza virus as key evidence. They also found many reverse mutations in the high-throughput sequencing data of patients to confirm this model. This important finding was reported, for the first time, to reveal the natural origin of SARS-CoV-2 at the molecular level.

The putative proteins (named as P1 to P17) were predicted from Nankai CDS , a 465- or 468-bp genomic region to name ORF 3b in SARS-CoV (AY274119: 25689-26153), and also ORF 3b in SARS-CoV-2 (MN908947: 25814-26281). Although the existence of these putative proteins have not been validated by any experiment (e.g. proteomics) and the Internal Ribosome Entry Sites (IRESs) of coronavirus are still unknown, this phenomena was still defined as alternative translation. The 17 putative proteins can be used for the genotyping of BB coronaviruses. The genotype of SARS-CoV-2 is P10P11P17P12P13. The Nankai group proposed two criteria to determine the

intermediate host(s): (1) the intermediate host(s) carry at least one strain of coronavirus with the FCS in the S protein. (2) the intermediate host(s) carry at least one strain of coronavirus with the genotype of P10P11P17P12P13.

---

## References

1. Jiayuan C, Shi J, Yau Tung O, Liu C, Li X, Zhao Q, Jishou R, Shan G. Bioinformatics Analysis of the 2019 Novel Coronavirus Genome[J]. Chinese Journal of Bioinformatics (In Chinese), 2020,18(2): 20-27.
2. Li X, Duan G, Zhang W, Shi J, Chen J, Chen S, Gao S, Ruan J. A Furin Cleavage Site Was Discovered in the S Protein of the 2019 Novel Coronavirus[J]. Chinese Journal of Bioinformatics (In Chinese), 2020,18(2): 28-33.
3. Duan G, Shi J, Xuan Y, Chen J, Liu C, Ruan J, Gao S, Li X. 5' UTR Barcode of the 2019 Novel Coronavirus Leads to Insights into Its Virulence[J]. Chinese Journal of Virology (In Chinese), 2020,36(2): 1-6.
4. Gao S, Duan G, Chen J, Wang L, Li X, Yau TO, Zhao Q, Zhou H, Xuan Y, Ruan J. A mutation model explaining acquisition of the furin cleavage site in the SARS-CoV-2 genome[J]. ResearchGate, 2020,2020(2020): 1-8.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/20405>