Deyin Guo ORCID iD: 0000-0002-8297-0814

Coronaviruses: genome structure, replication, and pathogenesis

Yu Chen[1, *], Qianyun Liu[1], Deyin Guo[2, *]

[1]State Key Laboratory of Virology, Modern Virology Research Center, College of Life Sciences, Wuhan University, Wuhan, P. R. China.

[2]Center for Infection & Immunity Study, School of Medicine, Sun Yat-sen University, Guangzhou, P. R. China.

*****Correspondence: chenyu@whu.edu.cn; guodeyin@mail.sysu.edu.cn

Abstract

The recent emergence of a novel coronavirus (2019-nCoV), which caused an outbreak of unusual viral pneumonia in tens of people in Wuhan, a central city of China, restated the risk of coronaviruses posed to public health. In this mini-review, we give a brief introduction of the general features of coronaviruses and describe various diseases caused by different coronaviruses in humans and animals. This review will help understand the biology and potential risk of coronaviruses that exist in richness in wildlife such as bats.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jmv.25681.

Introduction

Coronaviruses (CoVs) are important pathogens for human and vertebrates. They can infect respiratory, gastrointestinal, hepatic and central nervous system of human, livestock, avian, bat, mouse and many other wild animals [1-3]. Since the outbreaks of the severe acute respiratory syndrome (SARS) in 2002 and the Middle East respiratory syndrome (MERS) in 2012, the possibility of CoVs transmission from animals to human has been proved [4, 5]. Since the end of 2019, an outbreak of mystery pneumonia in Wuhan has been drawing tremendous attention around the world. Chinese government and researchers have taken swift measures to control the outbreak and conduct the etiological studies. The causative agent of the mystery pneumonia has been identified as a novel coronavirus by deep sequencing and etiological investigations by at least 5 independent laboratories of China. On 12 January 2020, the World Health Organization temporarily named the new virus as 2019 novel coronavirus (2019-nCoV). The formal name of the virus will be given by the International Committee of Taxonomy of Viruses (ICTV) according to the guidelines of viral nomenclature. The intermittent emergences and outbreaks of new types of coronaviruses remind us that CoVs are still a severe global health threat. Along with the changes of climate and ecology and the increased interaction of human and animals, new CoV outbreaks seem unavoidable in the future, and effective therapies and vaccines against CoVs must be developed as soon as possible.

Coronaviral, genome structure and replication

Coronaviruses belong to the subfamily *Coronavirinae* in the family of *Coronaviridae* of the order *Nidovirales*, and this subfamily includes four genera: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus* (Fig. 1A). The genome of CoVs is a single-stranded positive-sense RNA

(+ssRNA) (~30kb) with 5'-cap structure and 3'-poly-A tail. The genomic RNA is used as template to directly translate polyprotein (pp) 1a/1ab, which encodes non-structural proteins (nsps) to form replication-transcription complex (RTC) in a double-membrane vesicles (DMVs) [6]. And subsequently, a nested set of subgenomic RNAs (sgRNAs) are synthesized by RTC in a manner of discontinuous transcription [7]. These subgenomic mRNAs possess common 5′-leader and 3′-ends sequence. Transcription termination and subsequent acquisition of a leader RNA occurs at transcription regulatory sequences (TRS), located between ORFs. These minus-strand subgenomic RNAs serve as templates for the production of subgenomic mRNAs [8, 9].

Genomes and subgenomes of CoVs contain at least 6 open reading frames (ORFs). The first ORF (ORF1a/b), about two-third of genome length, encodes 16 non-structural proteins (nsp1-16), except *Gammacoronavirus* that lacks nsp1. There is a -1 frameshift between ORF1a and ORF1b, leading to production of two polypeptides: pp1a and pp1ab. These polypeptides will be processed by virally encoded chymotrypsin-like protease ($3CL^{pro}$) or main protease (M^{pro}) and one or two papain-like protease (PLPs) into 16 nsps [10,11]. Other ORFs on the one-third of genome near the 3' terminus encodes at least four main structural proteins: spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins. Besides these four main structural proteins, different CoVs encode special structural and accessory proteins, such as HE protein, 3a/b protein, 4a/b protein and *et al*. (Fig. 1B, lower panel). All the structural and accessory proteins are translated from the subgenomic RNAs of CoVs [7].

The genome alignment of CoVs shows 58% identity of nsp-coding region and 43% identity of structural proteins-coding region among different coronaviruses, with 54% at the whole genome level (Fig. 1B, upper panel), suggesting the

non-structural proteins are more conserved and the structural proteins have more diversity to fit varying environment. As the mutation rates in the replication of RNA viruses are much higher than that of DNA viruses, the genomes of RNA viruses are usually less than 10 K nucleotides. The genome size of CoV (~30kb) is the largest among all RNA viruses, which is almost two times larger than that of the second largest RNA viruses. The maintenance of the giant genome size of CoVs might be related to special features of the CoV RTC, which contains several RNA processing enzymes such as the 3'-5' exoribonuclease of nsp14. The 3'-5' exoribonuclease is unique to CoVs among all RNA viruses, and proved to function as a proofreading part of the RTC [12-14]. Sequence analysis showed that the 2019-nCoV possesses a typical genome structure of coronavirus and belongs to the cluster of *betacoronaviruses* that includes Bat-SARS-like (SL)-ZC45, Bat-SL ZXC21, SARS-CoV and MERS-CoV. Based on the phylogenetic tree of CoVs, 2019-nCov is more closely related to bat-SL-CoV ZC45 and bat-SL-CoV ZXC21 and more distantly related to SARS-CoV (Fig. 1A).

Functions of non-structural and structural proteins in coronaviral replication

Most of the non-structural proteins nsp1-16 have been reported for their specific roles in the replication of CoVs. However, the functions of some of nsps are unknown or not understood thoroughly. We summarize the known functions of the 16 nsps in Table. 1.

Four structural proteins are essential for virion assembly and infection of CoVs. Homotrimers of S proteins make up the spike on the surface of virus particles and it is the key for the viral attachment to host receptor [50, 51]. The M protein has three transmembrane domains and it shapes the virions, promotes membrane curvature, and binds to the nucleocapsid [52, 53]. The E protein plays a role in

virus assembly and release, and it is required for pathogenesis [54, 55]. The N protein contains two domains, both of them can bind virus RNA genome via different mechanisms. It is reported that N protein can bind nsp3 protein to help tether the genome to replicase-transcriptase complex (RTC), and package the encapsidated genome into virions [56-58]. N is also an antagonist of interferon and viral encoded repressor (VSR) of RNA interference (RNAi), which benefit the viral replication [59].

Diversity of coronavirus pathogenesis

Host range and tissue tropism show a lot of variation among different CoVs. Generally, the *alphacoronaviruses* and *betacoronaviruses* can infect mammals, and the *gammacoronaviruses* and *deltacoronaviruses* can infect birds, but some of them can also infect mammals [4, 60]. Before 2019, there were only six CoVs that can infect human and cause respiratory diseases: i) HCoV-229E, HCoV-OC43, HCoV-NL63 and HKU1 induce only mild upper respiratory disease, and in rare cases some of them can cause severe infection in infants, young and elders; ii) SARS-CoV and MERS-CoV can infect lower respiratory tract and cause severe respiratory syndrome in human [56, 61]. The new coronavirus 2019-nCoV, which belongs to *betacoronaviruses* according to the genome analysis (Fig. 1A), can also infect lower respiratory tract and cause pneumonia, but in general, the symptoms are milder than SARS and MERS. Until 18 January 2020, 62 infection cases were confirmed in Wuhan by sequence analysis, clinical diagnosis and epidemiological investigation, and among 62 cases, 8 cases showed severe symptoms. In addition, 2 cases were reported in Thailand and 1 case was reported in Japan; all these three cases had stayed in or visited Wuhan in two weeks before the onset of symptoms. Until now, two death cases have been recorded, and both of them are over 60 years old and have other diseases before the infection, one with abdominal tumor

and chronic liver disease and the other with myocarditis and renal dysfunction (http://wjw.wuhan.gov.cn). Not all cases have apparent connection with the Wuhan Huanan Seafood Wholesale Market that is believed to be original place of the outbreak of the 2019-nCoV. However, the animal source of the new virus is still not clear, and limited human-to-human transmission cannot be excluded.

Some CoVs can infect livestock, birds, bats, mice, whales and many other wild animals, and they can cause great economic loss. For example, in 2016, an HKU2-related bat coronavirus, swine acute diarrhea syndrome coronavirus (SADS-CoV), caused a large-scale outbreak of fatal disease in pigs in Southern China, and more than 24000 piglets dead [62]. This event is the first documented spillover of a bat coronavirus that caused severe disease in livestock [4, 63]. Although 2019-nCoV broke out in Hunan Seafood Market of Wuhan, the possibility of the 2019-nCoV transmission from seafood to human is unlikely. The 2019-nCoV and Beluga Whale CoV/SW1 belong to different genera and apparently have different host spectrum. As the Wuhan seafood market also sells other animals, the natural host of 2019-nCoV awaits to be identified. Due to the possibility of transmission from animal to human, CoVs in livestock and other animals including bats should be constantly monitored. We list the major pathogenic CoVs in Table. 2 for better understanding the pathogenesis of CoVs.

Treatment and prevention

At present, there are no specific antiviral therapies for coronavirus and the main treatments are supportive. Recombinant interferons (IFN) with ribavirin only have limited effects on coronaviruses infection [64]. After SARS and MERS epidemic, lots of anti-CoV agents have been developed against CoVs proteases, polymerases, MTases, and entry proteins, however none of them have been proved in clinical trials yet [65-67]. So far, the therapies with plasma and antibody

obtained from convalescent patients have been proposed as principal treatment [68].

There are various vaccine strategies for CoVs. Inactivated viruses, live-attenuated viruses, viral vector-based vaccines, subunit vaccines, recombinant proteins and DNA vaccines were developed but tested only in animals so far [69, 70].

As there is no effective therapeutics or vaccines, the best way to deal with severe infections of CoVs is to control the source of infection, early diagnosis, reporting, isolation, supportive treatments, and timely publishing epidemic information to avoid unnecessary panic. For individuals, good personal hygiene, fitted mask, ventilation and avoiding crowded places will help preventing CoVs infection.

Reference

1. Wang, L.F., et al., *Review of bats and SARS.* Emerg Infect Dis, 2006. **12**(12): p. 1834-40.
2. Ge, X.Y., et al., *Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor.* Nature, 2013. **503**(7477): p. 535-8.
3. Chen, Y. and D. Guo, *Molecular mechanisms of coronavirus RNA capping and methylation.* Virol Sin, 2016. **31**(1): p. 3-11.
4. Cui, J., F. Li, and Z.L. Shi, *Origin and evolution of pathogenic coronaviruses.* Nat Rev Microbiol, 2019. **17**(3): p. 181-192.
5. Cauchemez, S., et al., *Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and how to tell them apart.* Euro Surveill, 2013. **18**(24).
6. Snijder, E.J., et al., *Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex.* J Virol, 2006. **80**(12): p. 5927-40.
7. Hussain, S., et al., *Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus.* J Virol, 2005. **79**(9): p. 5288-95.
8. Sawicki, S.G., D.L. Sawicki, and S.G. Siddell, *A contemporary view of coronavirus transcription.* J Virol, 2007. **81**(1): p. 20-9.

9. Perlman, S. and J. Netland, *Coronaviruses post-SARS: update on replication and pathogenesis.* Nat Rev Microbiol, 2009. **7**(6): p. 439-50.

10. Masters, P.S., *The molecular biology of coronaviruses.* Adv Virus Res, 2006. **66**: p. 193-292.

11. Ziebuhr, J., E.J. Snijder, and A.E. Gorbalenya, *Virus-encoded proteinases and proteolytic processing in the Nidovirales.* J Gen Virol, 2000. **81**(Pt 4): p. 853-79.

12. Eckerle, L.D., et al., *Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing.* PLoS Pathog, 2010. **6**(5): p. e1000896.

13. Ogando, N.S., et al., *The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity.* Front Microbiol, 2019. **10**: p. 1813.

14. Smith, E.C., et al., *Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics.* PLoS Pathog, 2013. **9**(8): p. e1003565.

15. Huang, C., et al., *SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage.* PLoS Pathog, 2011. **7**(12): p. e1002433.

16. Tanaka, T., et al., *Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA.* J Virol, 2012. **86**(20): p. 11128-37.

17. Graham, R.L., et al., *The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication.* J Virol, 2005. **79**(21): p. 13399-411.

18. Gadlage, M.J., R.L. Graham, and M.R. Denison, *Murine coronaviruses encoding nsp2 at different genomic loci have altered replication, protein expression, and localization.* J Virol, 2008. **82**(23): p. 11964-9.

19. Lei, J., Y. Kusov, and R. Hilgenfeld, *Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein.* Antiviral Res, 2018. **149**: p. 58-74.

20. Serrano, P., et al., *Nuclear magnetic resonance structure of the nucleic acid-binding domain of severe acute respiratory syndrome coronavirus nonstructural protein 3.* J Virol, 2009. **83**(24): p. 12998-3008.

21. Beachboard, D.C., J.M. Anderson-Daniels, and M.R. Denison, *Mutations across murine hepatitis virus nsp4 alter virus fitness and membrane modifications.* J Virol, 2015. **89**(4): p. 2080-9.

22. Gadlage, M.J., et al., *Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function.* J Virol, 2010. **84**(1): p. 280-90.

23. Stobart, C.C., et al., *Chimeric exchange of coronavirus nsp5 proteases (3CLpro) identifies common and divergent regulatory determinants of protease activity.* J Virol, 2013. **87**(23): p. 12611-8.

24. Zhu, X., et al., *Porcine deltacoronavirus nsp5 inhibits interferon-beta production through the cleavage of NEMO.* Virology, 2017. **502**: p. 33-38.

25. Zhu, X., et al., *Porcine Deltacoronavirus nsp5 Antagonizes Type I Interferon Signaling by Cleaving STAT2.* J Virol, 2017. **91**(10).

26. Angelini, M.M., et al., *Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles.* mBio, 2013. **4**(4).

27. Cottam, E.M., M.C. Whelband, and T. Wileman, *Coronavirus NSP6 restricts autophagosome expansion.* Autophagy, 2014. **10**(8): p. 1426-41.

28. Kirchdoerfer, R.N. and A.B. Ward, *Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors.* Nat Commun, 2019. **10**(1): p. 2342.

29. Zhai, Y., et al., *Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer.* Nat Struct Mol Biol, 2005. **12**(11): p. 980-6.

30. te Velthuis, A.J., S.H. van den Worm, and E.J. Snijder, *The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension.* Nucleic Acids Res, 2012. **40**(4): p. 1737-47.

31. Egloff, M.P., et al., *The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique*

in the RNA virus world. Proc Natl Acad Sci U S A, 2004. **101**(11): p. 3792-6.

32. Zeng, Z., et al., *Dimerization of Coronavirus nsp9 with Diverse Modes Enhances Its Nucleic Acid Binding Affinity.* J Virol, 2018. **92**(17).

33. Bouvet, M., et al., *Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes.* J Biol Chem, 2014. **289**(37): p. 25783-96.

34. Chen, Y., et al., *Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex.* PLoS Pathog, 2011. **7**(10): p. e1002294.

35. Decroly, E., et al., *Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex.* PLoS Pathog, 2011. **7**(5): p. e1002059.

36. Ma, Y., et al., *Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex.* Proc Natl Acad Sci U S A, 2015. **112**(30): p. 9436-41.

37. Fang, S.G., et al., *Proteolytic processing of polyproteins 1a and 1ab between non-structural proteins 10 and 11/12 of Coronavirus infectious bronchitis virus is dispensable for viral replication in cultured cells.* Virology, 2008. **379**(2): p. 175-80.

38. Ahn, D.G., et al., *Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates.* Arch Virol, 2012. **157**(11): p. 2095-104.

39. te Velthuis, A.J., et al., *The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent.* Nucleic Acids Res, 2010. **38**(1): p. 203-14.

40. Adedeji, A.O. and H. Lazarus, *Biochemical Characterization of Middle East Respiratory Syndrome Coronavirus Helicase.* mSphere, 2016. **1**(5).

41. Hao, W., et al., *Crystal structure of Middle East respiratory syndrome coronavirus helicase.* PLoS Pathog, 2017. **13**(6): p. e1006474.

42. Jia, Z., et al., *Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis.* Nucleic Acids Res, 2019. **47**(12): p. 6538-6550.

43. Bouvet, M., et al., *RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex.* Proc Natl Acad Sci U S A, 2012. **109**(24): p. 9372-7.

44. Chen, Y., et al., *Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase.* Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3484-9.

45. Minskaia, E., et al., *Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis.* Proc Natl Acad Sci U S A, 2006. **103**(13): p. 5108-13.

46. Bhardwaj, K., et al., *RNA recognition and cleavage by the SARS coronavirus endoribonuclease.* J Mol Biol, 2006. **361**(2): p. 243-56.

47. Deng, X., et al., *Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages.* Proc Natl Acad Sci U S A, 2017. **114**(21): p. E4251-E4260.

48. Zhang, L., et al., *Structural and Biochemical Characterization of Endoribonuclease Nsp15 Encoded by Middle East Respiratory Syndrome Coronavirus.* J Virol, 2018. **92**(22).

49. Shi, P., et al., *PEDV nsp16 negatively regulates innate immunity to promote viral proliferation.* Virus Res, 2019. **265**: p. 57-66.

50. Beniac, D.R., et al., *Architecture of the SARS coronavirus prefusion spike.* Nat Struct Mol Biol, 2006. **13**(8): p. 751-2.

51. Delmas, B. and H. Laude, *Assembly of coronavirus spike protein into trimers and its role in epitope expression.* J Virol, 1990. **64**(11): p. 5367-75.

52. Nal, B., et al., *Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E.* J Gen Virol, 2005. **86**(Pt 5): p. 1423-34.

53. Neuman, B.W., et al., *A structural analysis of M protein in coronavirus assembly and morphology.* J Struct Biol, 2011. **174**(1): p. 11-22.

54. DeDiego, M.L., et al., *A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo.* J Virol, 2007. **81**(4): p. 1701-13.

55. Nieto-Torres, J.L., et al., *Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis.* PLoS Pathog, 2014. **10**(5): p. e1004077.

56. Fehr, A.R. and S. Perlman, *Coronaviruses: an overview of their replication and pathogenesis.* Methods Mol Biol, 2015. **1282**: p. 1-23.

57. Chang, C.K., et al., *Modular organization of SARS coronavirus nucleocapsid protein.* J Biomed Sci, 2006. **13**(1): p. 59-72.

58. Hurst, K.R., C.A. Koetzner, and P.S. Masters, *Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein.* J Virol, 2009. **83**(14): p. 7221-34.

59. Cui, L., et al., *The Nucleocapsid Protein of Coronaviruses Acts as a Viral Suppressor of RNA Silencing in Mammalian Cells.* J Virol, 2015. **89**(17): p. 9029-43.

60. Woo, P.C., et al., *Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus.* J Virol, 2012. **86**(7): p. 3995-4008.

61. Su, S., et al., *Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses.* Trends Microbiol, 2016. **24**(6): p. 490-502.

62. Zhou, P., et al., *Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin.* Nature, 2018. **556**(7700): p. 255-258.

63. Simas, P.V., et al., *Bat coronavirus in Brazil related to appalachian ridge and porcine epidemic diarrhea viruses.* Emerg Infect Dis, 2015. **21**(4): p. 729-31.

64. Cinatl, J., et al., *Treatment of SARS with human interferons.* Lancet, 2003. **362**(9380): p. 293-4.

65. Chan, J.F., et al., *Broad-spectrum antivirals for the emerging Middle East respiratory syndrome coronavirus.* J Infect, 2013. **67**(6): p. 606-16.

66. Cheng, K.W., et al., *Thiopurine analogs and mycophenolic acid synergistically inhibit the papain-like protease of Middle East respiratory syndrome coronavirus.* Antiviral Res, 2015. **115**: p. 9-16.

67. Wang, Y., et al., *Coronavirus nsp10/nsp16 Methyltransferase Can Be Targeted by nsp10-Derived Peptide In Vitro and In Vivo To Reduce Replication and Pathogenesis.* J Virol, 2015. **89**(16): p. 8416-27.

68. Mair-Jenkins, J., et al., *The effectiveness of convalescent plasma and hyperimmune immunoglobulin for the treatment of severe acute respiratory infections of viral etiology: a systematic review and exploratory meta-analysis.* J Infect Dis, 2015. **211**(1): p. 80-90.

69. Graham, R.L., E.F. Donaldson, and R.S. Baric, *A decade after SARS: strategies for controlling emerging coronaviruses.* Nat Rev Microbiol, 2013. **11**(12): p. 836-48.

70. de Wit, E., et al., *SARS and MERS: recent insights into emerging coronaviruses.* Nat Rev Microbiol, 2016. **14**(8): p. 523-34.

Figure

Figure 1. The genomic structure and phylogenetic tree of coronaviruses. (A) The phylogenetic tree of representative CoVs, with the new coronavirus 2019-nCoV highlighted in red. (B) The genome structure of four genera of coronaviruses. Pp1a and pp1b represent the two long polypeptides that are processed into 16 non-structural proteins. S, E, M and N indicate the four structural proteins spike, envelope, membrane and nucleocapsid. HE, hemagglutinin-esterase. Viral names: HCoV, human coronavirus; TGEV, transmissible gastroenteritis virus; MHV, murine hepatitis virus; HKU, coronaviruses identified by Hong Kong University; IBV, infectious bronchitis virus;

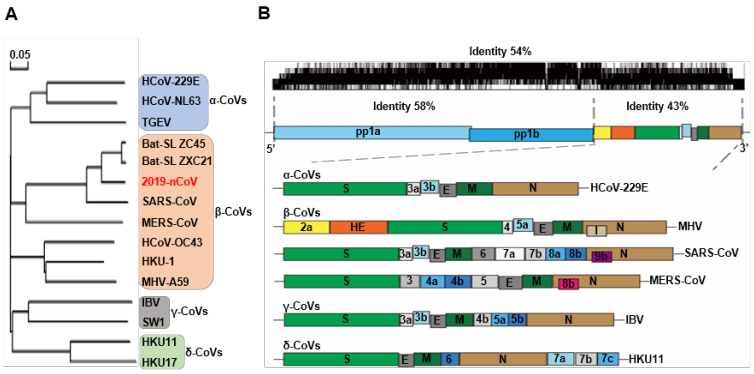

Table. 1 The 16 non-structural proteins of coronaviruses and their functions.

nsps	Functions	Reference
nsp1	cellular mRNA degradation, inhibiting IFN signaling	[15, 16]
nsp2	unknown	[17, 18]
nsp3	PLP, polypeptides cleaving, blocking host innate immune response, promoting cytokine expression	[19, 20]
nsp4	DMV formation	[21, 22]
nsp5	$3CL^{pro}$, M^{pro}, polypeptides cleaving, inhibiting IFN signaling	[23-25]
nsp6	restricting autophagosome expansion, DMV formation	[26, 27]
nsp7	cofactor with nsp8 and nsp12	[28, 29]
nsp8	cofactor with nsp7 and nsp12, primase	[28-30]
nsp9	dimerization and RNA binding	[31, 32]
nsp10	scaffold protein for nsp14 and nsp16	[33-36]
nsp11	unknown	[37]

nsp12	primer dependent RdRp	[28, 38, 39]
nsp13	RNA helicase, 5’ triphosphatase	[40-42]
nsp14	exoribonuclease, N7-MTase	[12, 43-45]
nsp15	endoribonuclease, evasion of dsRNA sensors	[46-48]
nsp16	2’-O-MTase; avoiding MDA5 recognition, negatively regulating innate immunity	[34, 35, 49]

Table 2. List of part of important pathogenic coronaviruses.

Virus	**Genus**	**Host**	**Symptoms**
Human CoV-229E	alpha	human	mild respiratory tract infections
Human CoV-NL63	alpha	human	mild respiratory tract infections
PRCV/ISU-1	alpha	pig	mild respiratory tract infections
TGEV/PUR46-MAD	alpha	pig	diarrhea, with 100% mortality in piglets less than 2weeks old

PEDV/ZJU-G1-2013	alpha	pig	severe watery diarrhea
SeACoV-CH/GD-01	alpha	pig	severe and acute diarrhoea and acute vomiting
Canine CoV/TU336/F/2008	alpha	dog	mild clinical signs, diarrhoea
Camel alphacoronavirus isolate camel/Riyadh	alpha	camel	asymptomatic
Feline infectious peritonitis virus	alpha	cat	fever, vasculitis, and serositis, with or without effusions
Human CoV-HKU1	beta	human	pneumonia
SARS-CoV	beta	human	severe acute respiratory syndrome, 10% mortality rate
MERS-CoV	beta	human	severe acute respiratory syndrome, 37% mortality rate
Bovine CoV/ENT	beta	cow	diarrhoea
Equine CoV/Obihiro12-1	beta	Horse	fever, anorexia, leucopenia

MHV-A59	beta	mouse	acute pneumonia and severe lung injuries
Beluga Whale CoV/SW1	gamma	whale	pulmonary disease, terminal acute liver failure
IBV	gamma	chicken	severe respiratory disease
Bulbul coronavirus HKU11	delta	bulbul	respiratory disease (collected from respiratory tract of dead wild birds)
Sparrow coronavirus HKU17	delta	bird	respiratory disease (collected from respiratory tract of dead wild birds)