

# Coronaviruses: Basic Virology, Nomenclature, Different Strains, Their Origin and Relationship, Host Specificity

**Xuming Zhang**

**UAMS**

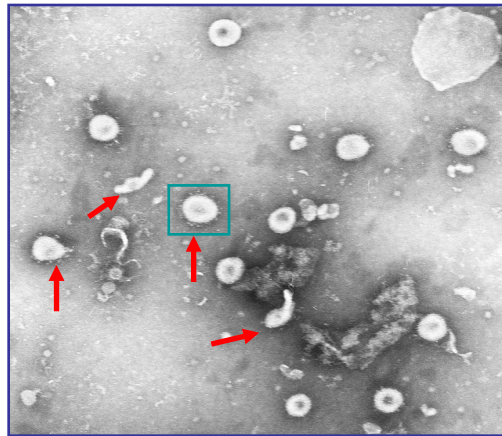
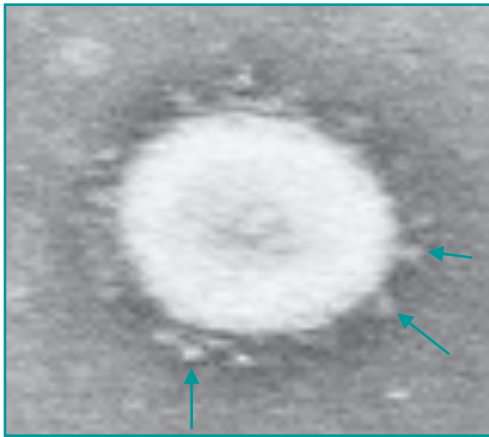


COLLEGE OF MEDICINE  
DEPARTMENT OF  
MICROBIOLOGY & IMMUNOLOGY

UNIVERSITY OF ARKANSAS FOR MEDICAL SCIENCES

# What Is Coronavirus?

-**Coronaviruses** are groups of viruses that have similar morphological features. The Latin name **Corona** is originally derived from its outlook; the virus particle (virion) looks like a **crow**n under an electron microscope (left panel). The virion is  $\approx 120\text{-}160$  nm in diameter and more or less spherical. However, it is important to note that the spherical shape is often observed for cell cultured viruses (that is what you have often seen in the textbook or internet) whereas viruses isolated from clinical samples are more pleomorphic (right panel).



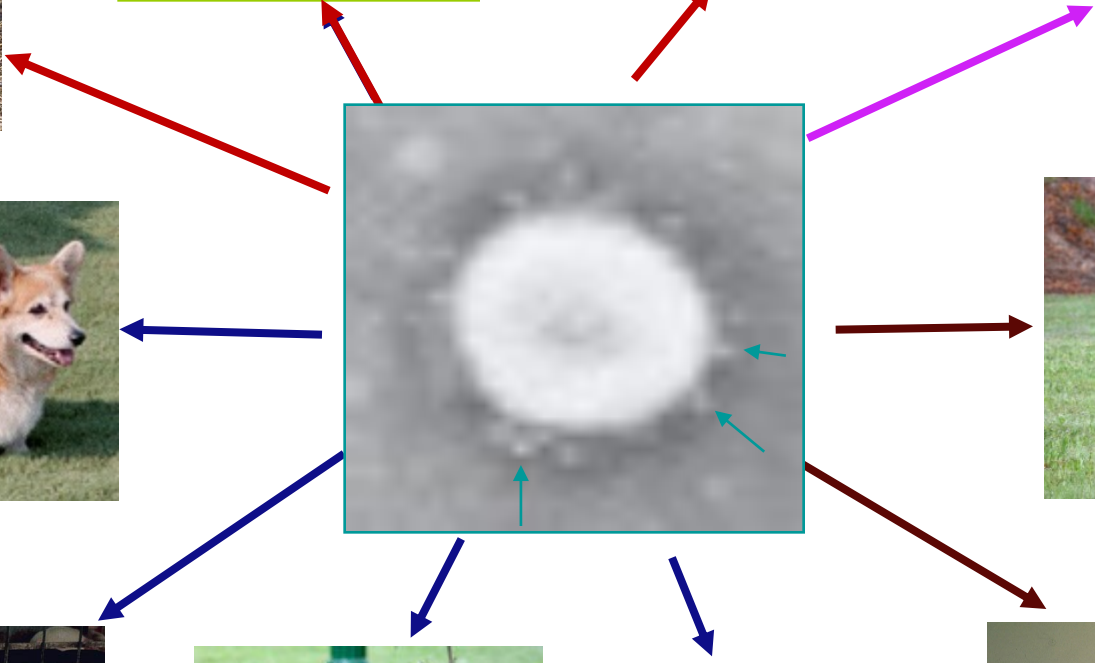
This is the human enteric coronavirus (HECoV) I isolated from fecal samples of a diarrheic child in early 1990's (Zhang et al., 1994. *J. Med. Virol.* 44:152-161). The picture was taken directly from EM in the fecal sample prior to cell culture.

# Nomenclature and Classification of the Members of Coronaviridae

-**Coronavirus** belongs to the family **Coronaviridae**, which is divided into two subfamilies: Coronavirinae and Torovirinae. Within Coronavirinae, there are four genera ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ). They can infect humans and various species of animals (see next 2 slides), causing diseases ranging from respiratory, digestive, neurological and immune-mediated diseases. The most common coronaviruses are listed in the table.

-In general, coronaviruses are species-specific, e.g. porcine coronavirus only infects pig; murine coronavirus only infects rodents, and so forth.

-However, in 1994, I was the first that described that the coronavirus (HECoV-4408) isolated from a diarrheic child is more closely related to bovine coronavirus than all known human coronaviruses at that time, suggesting a potential animal to human jumping. This **Zoonotic** phenomenon was subsequently confirmed during the first SARS outbreak in 2002-2003 and likely for the current COVID-19 outbreak as well. (Zhang et al., 1994. *J. Med. Virol.* 44:152-161).



# Coronaviridae

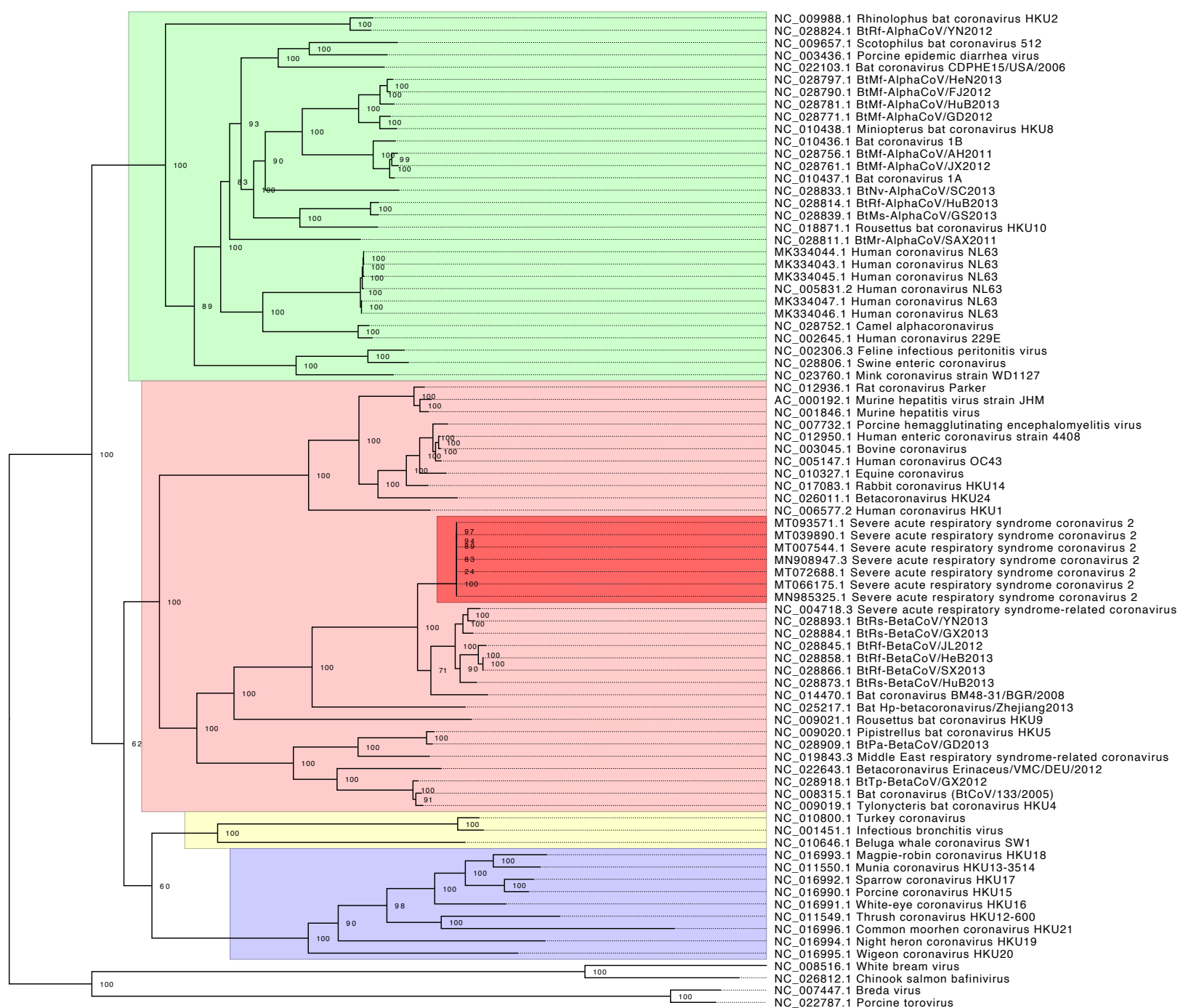
Subfam.	Genus	Representative members	Host and common diseases
Corona- virinae	Alpha	Human coronavirus 229E (HCoV-229E)	Human, common cold
		Transmissible gastroenteritis virus (TGEV)	Pig, gastroenteritis
		Porcine epidemic diarrhea virus (PEDV)	Pig, gastroenteritis
		Feline infectious peritonitis virus (FIPV)	Cat, peritonitis, immune-dis.
		Feline enteric coronavirus (FECoV)	Cat, enteritis
		Canine coronavirus virus (CCoV)	Dog, enteritis
		Human coronavirus NL63 (HCoV-NL63)	Human, respiratory disease
	Beta	Mouse hepatitis virus (MHV)	Rodent, hepatitis, CNS disease
		Porcine hemaggl.encephalomyetitis virus	Pig, CNS disease
		Bovine coronavirus (BCoV)	Cattle, GI/respiratory diseases
		Human coronavirus (HCoV)-OC43	Human, common cold
		Human enteric coronavirus (HECoV)-4408	Human, diarrhea
		Human coronavirus HKU1(HCoV-HKU1)	Human, respiratory diseases
		Severe acute respiratory syndrome(SARS)-CoV	Human, SARS
Mid-East respiratory syndrome (MERS)-CoV	Human, SARS		
SARS-CoV-2	Human, SARS (COVID-19)		
Bat coronavirus (Bat-CoV)	Bat, disease unknown		
Gamma	Avian Infectious bronchitis virus (IBV)	Chicken, bronchitis	
	Turkey enteric coronavirus (TECoV)	Turkey, enteritis	
Delta	Porcine deltacoronavirus (PdCoV)	Pig, enteritis	
Toro- virinae	Berne virus (BV)	Horse, enteritis	
	Breda virus	Cattle, enteritis	



# Origin and Phylogenetic Relationship Among Various Coronaviruses

- Phylogenetic analysis shows that there are at least 3 subgroups or clades within the genus Betacoronavirus (see next slide for phylogenetic tree). SARS-CoV-2 (causing the current COVID-19) is most closely related to SARS-CoV and bat coronaviruses isolated from China. In fact, it is more closely related to some of the bat coronaviruses than to SARS-CoV, suggesting that SARS-CoV-2 is likely originated from bat coronavirus.
- However, currently there is no evidence suggesting that SARS-CoV-2 is directly transmitted from bat to humans. It is more likely through intermediate carriers, such as pangolin.
- Previous (2002-2003) outbreak of SARS-CoV was traced back to wild animals, such as racoon dog and civet cat, as intermediate carriers. Subsequent large-scale surveillance isolated a large number of coronaviruses from bats and sequence data support the notion that SARS-CoV was originated from bats as seen in the phylogenetic tree.
- MERS-CoV also appears to originate in bat and is transmitted through camels to humans.

# Phylogenetic relatedness of coronaviruses

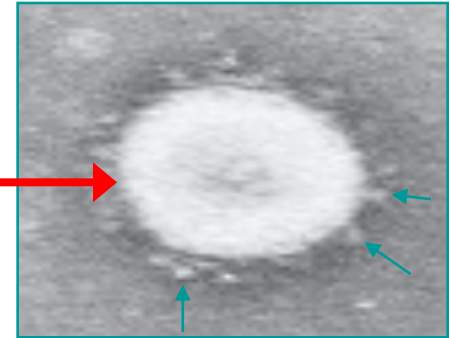


0.4

Courtesy of Visanu Wachai, unpublished data.

# What are the structural components of a Coronavirus?

-**Coronavirus** is an enveloped RNA virus, which means that the virion has a bilayer lipid membrane (derived from host cell membrane), an “**Envelope**”, wrapping the RNA genome inside the nucleocapsid.



-Clinical relevance of the viral envelope: During the current COVID-19 outbreak, there are a huge number of the public, asking the health professionals what kinds of disinfectants can be used to effectively kill coronavirus, Lysol, Clorox, Alcohol, and soaps. To answer this question, it is important to understand the biophysics of the envelope. Unlike naked (non-enveloped) viruses, the presence of an envelope on the virion makes coronavirus relatively easier to be “killed” by heat and detergents. Any detergent containing chemicals that can break down or “melt” the lipid membrane can render coronaviruses noninfectious (dead), because the spike protein (see next slide) that recognizes and binds the receptor of the host cells (i.e. nasal epithelium) for infection requires the envelope to be anchored on the virion surface. Without the envelope, the virus is essentially dead.



-There are 4 to 5 proteins associated with coronavirus Virion depending on virus strains:

**-Spike (S) protein:**

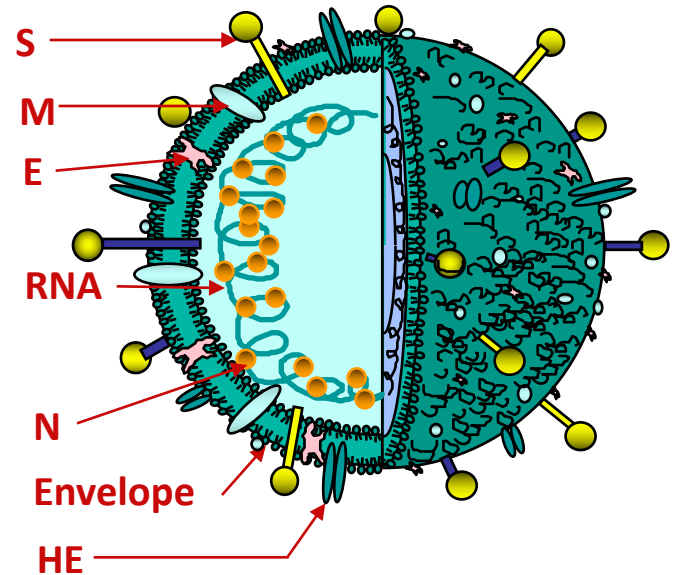
-The spike protein is the most important component of coronavirus for infection.

-Its amino terminal domain (S1) forms the typical globular shape (Crown) protruding from the virion surface while its carboxyl terminus (S2) forms the stem that anchors the spike to the envelope.

-Within the globular domain, there is a stretch of amino acids that form the Receptor-Binding Domain (RBD), which recognizes a specific receptor on the host cells (a target for drug and vaccine).

-It facilitates fusion between viral envelope and cell membrane during infection (entry) (a target for drug).

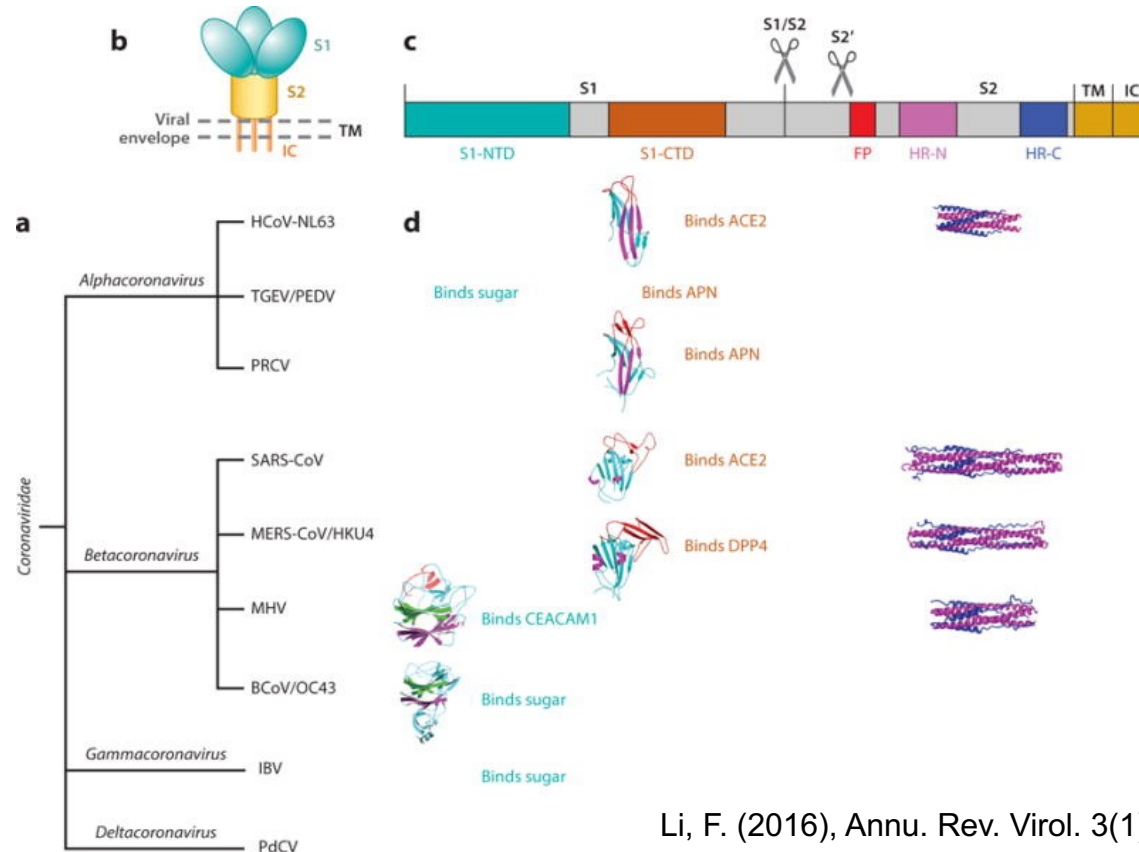
-it also elicits the production of neutralizing antibodies and cell-mediated immunity (a target for vaccine).



# -Clinical Relevance of the Spike Protein:

## 1. Host species specificity, animal to human transmission, human to human spread.

-The species-specificity of coronavirus infection is largely determined by the interaction between the spike protein and its host receptor.



Li, F. (2016), Annu. Rev. Virol. 3(1):237-261.

The spike proteins (b,c) of various coronaviruses (a) recognize their respective species-specific receptors (d).

(Continued)

- Interestingly, 3 human coronaviruses (HCoV-NL63, SARS-CoV, and SARS-CoV-2) utilize the same angiotensin-converting enzyme-2 (ACE2) as a receptor for virus infection.
- Transmembrane serine protease 2 (TMPRSS2) has also been shown as an additional receptor that enhances the binding of Spike to ACE2 during SARS-Cov-2 infection.
- It was initially speculated that the higher binding affinity of the spike protein of SARS-CoV-2 to ACE2 receptor might be one of the factors contributing to the higher transmissibility during COVID-19 than SARS-CoV during the previous outbreak.
- Subsequently, numerous Variants arisen during the course of the pandemic lend support of this hypothesis, with some Variants being more pathogenic and lethal.
- From evolutionary standpoint, mutations in RNA viruses in general and in the attachment proteins in particular are very common and of high frequency.

(Continued)

-Let me give you an example for better understanding the mutation frequency in relationship to the current COVID pandemic.

-For example, the mutation rate for a single-stranded RNA phage Q $\beta$  is about  $1.5 \times 10^{-3}$  mut/nt/rep. Based on this rate, there would be  $\approx 45$  nucleotide (nt) mutations in SARS-CoV-2 genome ( $\approx 30$  kb in length) following a single replication cycle. Assuming the virus replicates 3 times per day for 7 days in a patient, then there would be  $\approx 1,000$  nt mutations per genome at the end. That might seem unfathomable.

-Fortunately, we do not see so many mutations in coronavirus genome. If you compare the sequence of the virus isolated today with that of the first isolate in 2019, you would not see thousands of nt differences.

-Unlike the error-prone nature of the RNA polymerases in other RNA viruses, coronavirus RNA polymerase complex has some kind of “proofreading” activity, resulting in less frequent mutations. This is why you only see a limited number of Variants circulating in the populations throughout the world during the entire pandemic period.

-It is expected that, with increasing vaccination, antibodies produced in the vaccinated population will drive more mutations in the spike gene due to evolutionary pressure.

(Continued)

**2. The spike protein is a target for antiviral drug development.**

-Because the spike protein exerts two functions, receptor-binding and fusion, both of which are essential for virus infection, drugs that block its receptor-binding and fusion would effectively inhibit virus infection.

**3. The spike protein is the most important target for vaccine development.**

-Because the spike protein elicits neutralizing antibodies and cell-mediated immunity, antigens that contain the spike would be good candidates for COVID-19 vaccines. In fact, any effective COVID vaccines must include the spike.



## Hemagglutinin/esterase (HE) protein:

- Only present in certain Betacoronaviruses, SARS-CoVs do not have the HE protein.
- It has receptor-binding (hemagglutination) and receptor-destroying (acetyl-esterase) activities. Binds to 9-O-acetylneuraminic acid-containing receptor, similar to influenza C virus HEF protein. We suggested that it was acquired from influenza C virus through recombination events (Zhang, Kousoulas, and Storz.1992. Virology 186:318-323)
- It is not essential for virus infection but may be involved in viral pathogenesis (Zhang et al.1998. Virology 242:170-183).

## Membrane (M) protein:

- A transmembrane protein embedded in the envelope.
- Essential for virion assembly.

## Envelope (E) protein:

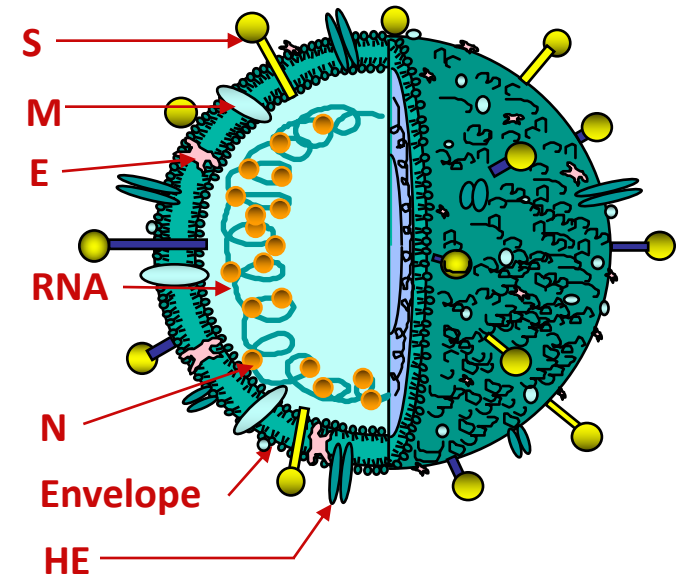
- Small envelope protein embedded in the envelope.
- Critical for virion assembly.

## Nucleocapsid (N) protein:

- Interacting with viral RNA genome to form nucleocapsid.
- Involved in encapsidation and virion assembly.
- Possibly also in regulation of replication and transcription.

## The RNA genome:

- The viral genome is a single-strand, positive-sense RNA of 30-32 kb in length, the largest genome among RNA viruses.
- It has a cap at the 5'-end and a poly(A) tail at the 3'-end, a typical mRNA.
- It contains multiple ORFs (genes). However, only the 5'-most ORF (gene 1ab) is translated from the genome, which encodes viral RNA-dependent RNA polymerase and related proteins(ns1-16).



Genome organization of SARS-CoV

# Coronavirus Life Cycle

-Coronavirus life cycle begins with attachment of viral spike protein to the receptor on cell surface, which then either induces fusion between viral envelope and plasma membrane and delivers the nucleocapsid into cytoplasm, or mediates endocytosis and induces fusion between viral envelope and endosomal membrane (**chloroquine blocks virus nonspecifically from releasing from endosome**)

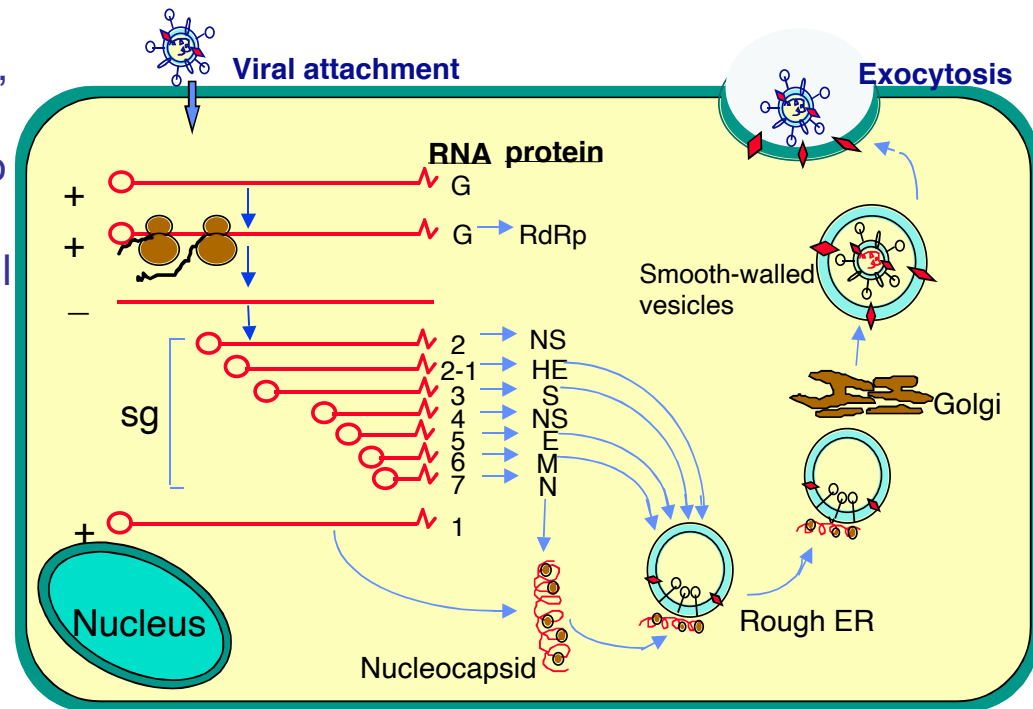
-Upon entry, the viral genomic RNA serves as an mRNA for translation of the 5'-most ORF (gene 1ab), which is then processed into 16 proteins (ns1-16), many of which are involved in viral replication, including the RNA-dependent RNA polymerase (**important targets for antiviral drugs**).

-Unlike any other RNA viruses, coronavirus synthesizes, in addition to the genomic RNA, a nested set of subgenomic (sg) mRNAs, each of which has a unique gene exposed to the 5'-end. This is indeed a clever way for the virus to overcome the inability of host cell translation machinery to express downstream ORFs from the genome.

-All viral structural proteins (the building blocks of virion) are then translated from the sg mRNAs.

-These structural proteins along with the genomic RNA are assembled into virions in intracellular vesicles and released from cell surface via exocytosis.

-In the respiratory tract, SARS-CoV-2 is released from the apical surface of infected epithelia. That is why the virus is so easily transmitted through droplets (large amount of virus) by coughing and sneezing.



# Conclusion Remark

In this introductory lecture, I have provided you with a glimpse of the basic biology of Coronaviruses without much mechanistic detail. Understanding these biological features is important for understanding the basis of the epidemiology, transmission, pathogenesis, antiviral drugs and vaccine development that will be discussed in a series of lectures that follow.