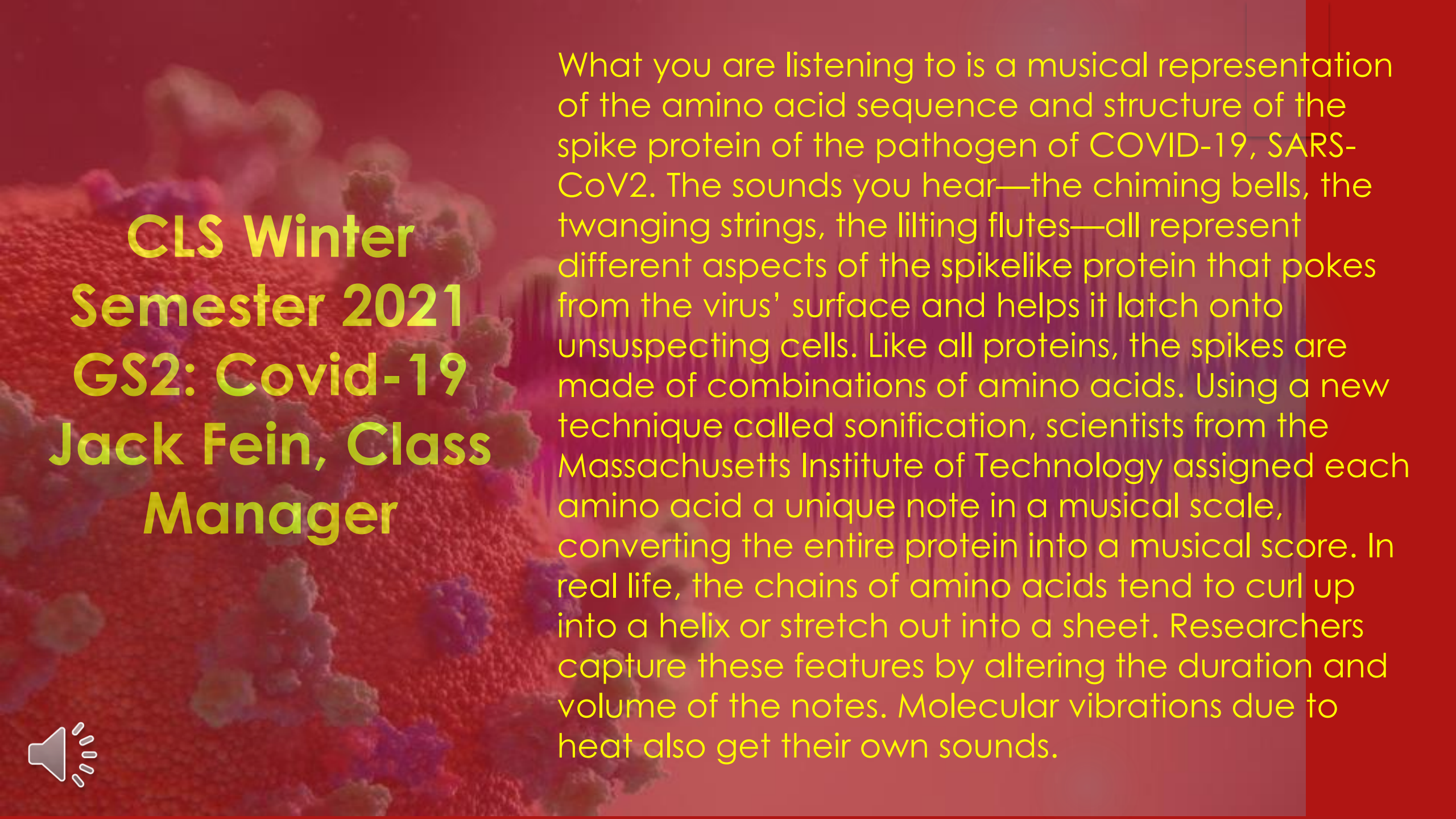




The Biology of Coronavirus

SARS-COVID-19

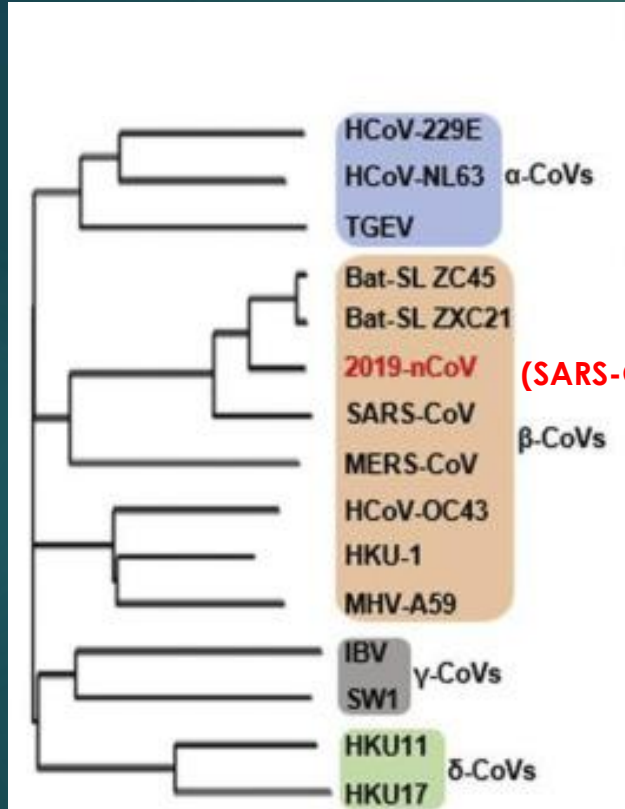


**CLS Winter
Semester 2021
GS2: Covid-19
Jack Fein, Class
Manager**

What you are listening to is a musical representation of the amino acid sequence and structure of the spike protein of the pathogen of COVID-19, SARS-CoV2. The sounds you hear—the chiming bells, the twanging strings, the lilting flutes—all represent different aspects of the spikelike protein that pokes from the virus' surface and helps it latch onto unsuspecting cells. Like all proteins, the spikes are made of combinations of amino acids. Using a new technique called sonification, scientists from the Massachusetts Institute of Technology assigned each amino acid a unique note in a musical scale, converting the entire protein into a musical score. In real life, the chains of amino acids tend to curl up into a helix or stretch out into a sheet. Researchers capture these features by altering the duration and volume of the notes. Molecular vibrations due to heat also get their own sounds.



Cladistics and Taxonomy of Coronaviruses



Order Nidovirales

Family: Coronaviridae

Subfamilies:

Alphacoronavirus (infect mammals)

Betacoronavirus (mammals)

Gammacoronavirus (birds and fish, occasionally mammals)

Deltacoronavirus (birds and fish, occasionally mammals)

Key:

HCoV: human coronaviruses

TGEV: transmissible gastroenteritis virus

Bat-SL: bat SARS-like coronavirus

SARS: Sudden Acute Respiratory Syndrome virus

MERS: Middle Eastern Respiratory Syndrome virus

MHV: Murine Hepatitis virus

IBV: Infectious Bronchitis virus

HKU: coronaviruses identified by Hong Kong University

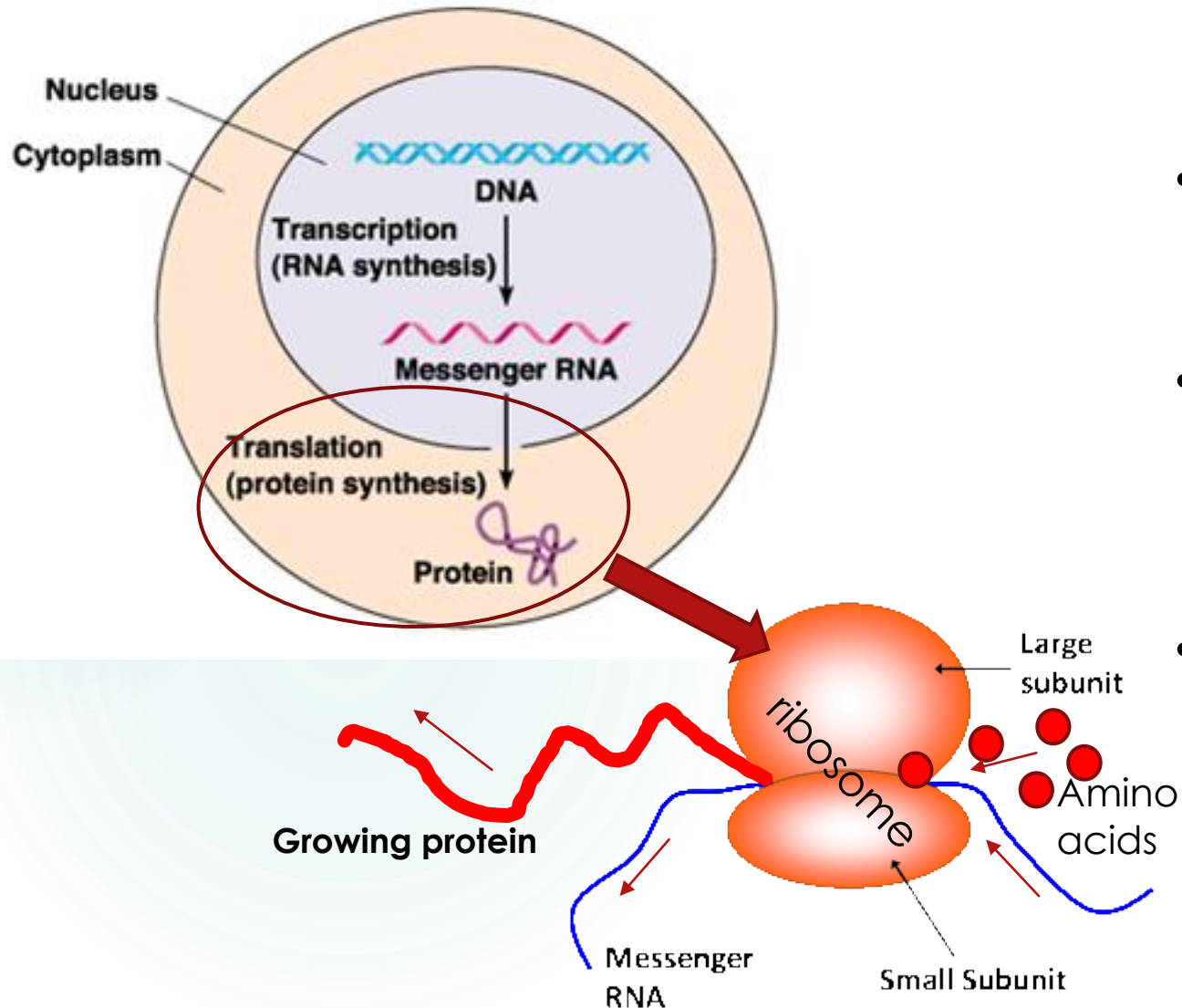
Coronaviruses and Disease

- ▶ Prior to 2019, only six coronaviruses were known to affect humans:
 - ▶ HCoV-229E, HCoVOC43, HCoVNL63
 - ▶ cause mild upper respiratory infections, though occasionally severe in infants and elders.
 - ▶ cause about 20% of “common colds” (the rest are mostly rhinoviruses)
 - ▶ HKU1
 - ▶ Causes pneumonia
 - ▶ SARS-CoV and MERS-CoV
 - ▶ Cause severe lower respiratory infections in humans (10% and 37% mortality, respectively)
- ▶ In 2019 a new coronavirus, 2019-nCoV (**n** for **novel**) arose in Wuhan, China, and spread worldwide, causing COVID-19 “Coronavirus Disease 2019”.
- ▶ 2019-nCoV was renamed to SARS-CoV2 in February 2020 by the by the Coronavirus Study Group of the International Committee on Taxonomy of Viruses based on its phylogenetic relationship to SARS-CoV.

What is a Virus?

- ▶ Viruses are usually made up of some genetic material (RNA or DNA) surrounded by a capsule commonly made of a lipid bilayer containing proteins.
- ▶ Unlike cells, viruses do not contain the machinery for reproducing themselves. In order to reproduce they must somehow invade a cell that has the necessary machinery and then take over that machinery and cause it to make new virus materials.
- ▶ Typically, this results in the death of the cell.

How Viruses Work

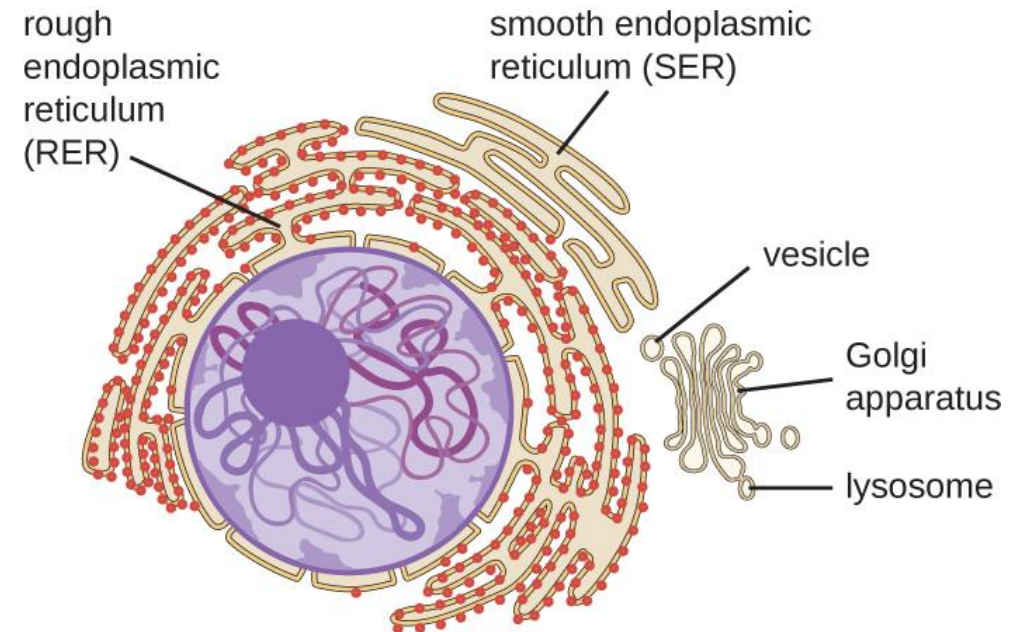
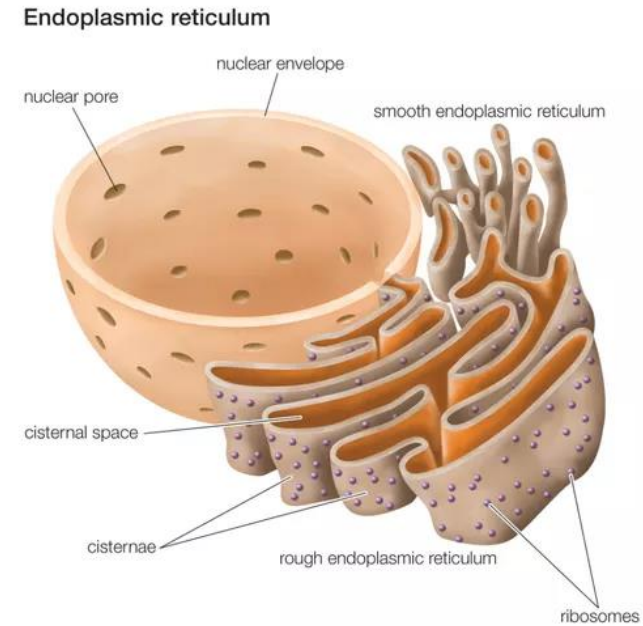


- Normal cells have machinery for making new materials in order to grow or reproduce or to replace worn out materials.
- The blueprint for these materials is found as genes, which are sections of DNA, kept protected in the nucleus.
- The information contained in the DNA is transcribed and transported out of the nucleus by an RNA copy of the gene, called messenger RNA (mRNA).
- The messenger RNA attaches to a ribosome, providing the ribosome with the information it needs.

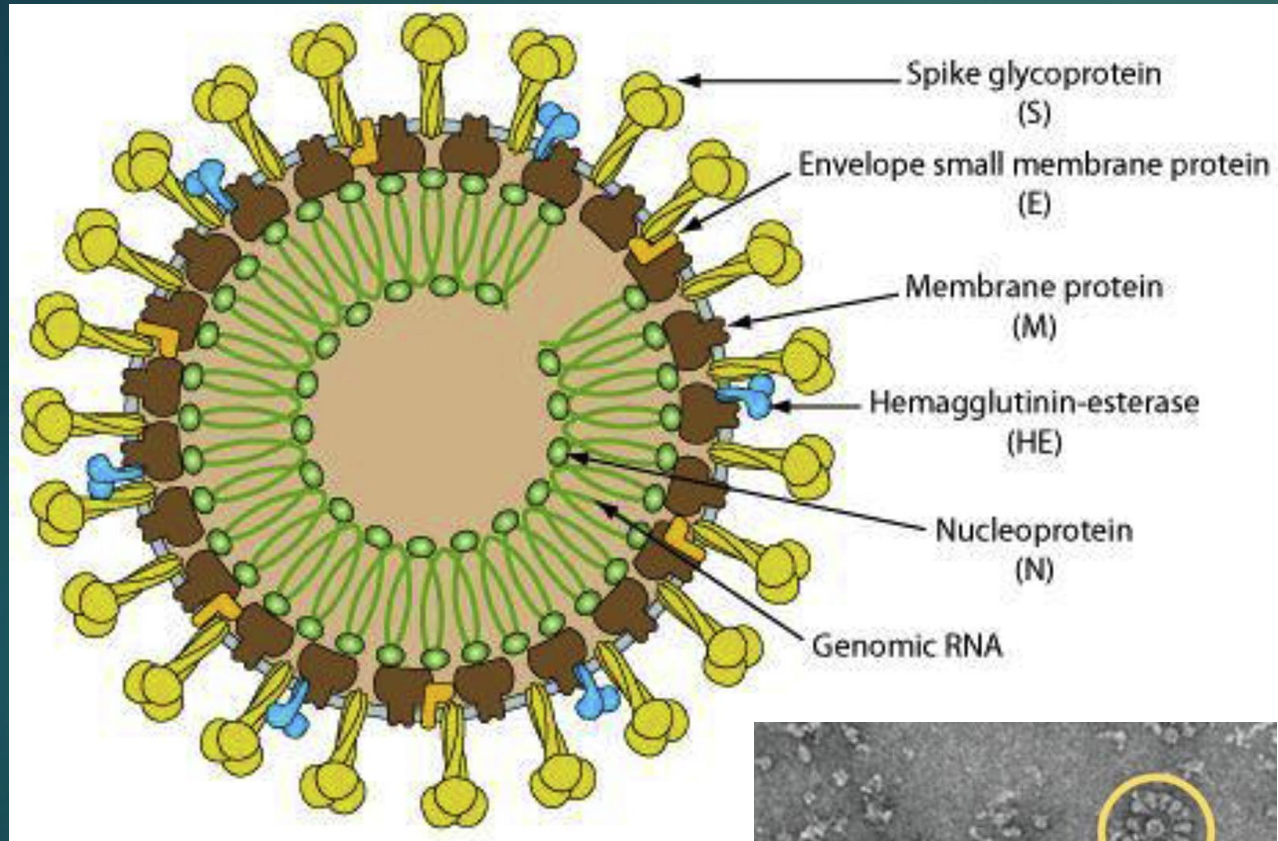
Coronaviruses bring their own messenger RNA into the cell and thus usurp the cell's machinery to produce their own viral proteins.

Formation of proteins in a cell: The Endoplasmic Reticulum

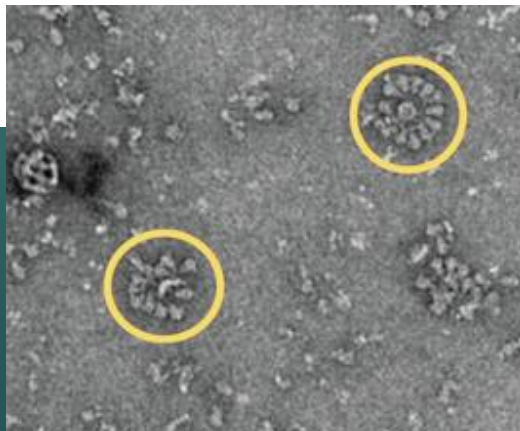
- ▶ Proteins designated for export as well as lipid bilayer membranes are manufactured in the ER.
- ▶ Ribosomes fabricate such proteins directly into the lumen of the RER.
- ▶ The proteins are finalized in the Golgi apparatus, wedged to membranes and then exported or emplaced in the cell's own membrane.



Anatomy of a killer



Spike glycoprotein



Genomic RNA:

- Single-stranded positive sense RNA
 - Largest genome of all RNA viruses (26-32 kb vs 13 kb for Influenza A (the 1918 Spanish Flu virus))
 - Has code for six viral structural proteins:
 - S (Spike)
 - E (Envelope)
 - M (Membrane)
 - N (Nucleoprotein or Nucleocapsid)
 - HE (Hemagglutinin-esterase)
 - 7a
 - These proteins have multiple roles, including interacting with the host cell proteins to advance the virus' reproduction

Non-Structural Protein Genes of SARS-CoV-19

- ▶ Nsp1: Inhibits host cell protein synthesis by cellular mRNA degradation, inhibiting IFN signaling
- ▶ Nsp2: Unknown
- ▶ Nsp3: PLP, polypeptides cleaving, blocking host innate immune response, promoting cytokine expression
- ▶ Nsp4: DMV formation
- ▶ Nsp5: 3CLpro, Mpro, polypeptides cleaving, inhibiting IFN signaling
- ▶ Nsp6: Restricting autophagosome expansion, DMV formation
- ▶ Nsp7: Cofactor with nsp8 and nsp12
- ▶ Nsp8: Cofactor with nsp7 and nsp12, primase
- ▶ Nsp9: Dimerization and RNA binding
- ▶ Nsp10: Scaffold protein for nsp14 and nsp16
- ▶ Nsp11: Unknown
- ▶ Nsp12: RdRp (RNA-dependent RNA polymerase)
- ▶ Nsp13: RNA helicase, 5' triphosphatase
- ▶ Nsp14: Exoribonuclease, N7-Mtase. **Involved in proofreading RNA replication**
- ▶ Nsp15: Endoribonuclease, evasion of dsRNA sensors
- ▶ Nsp16: 2'-O-MTase; avoiding MDA5 recognition, negatively regulating innate immunity

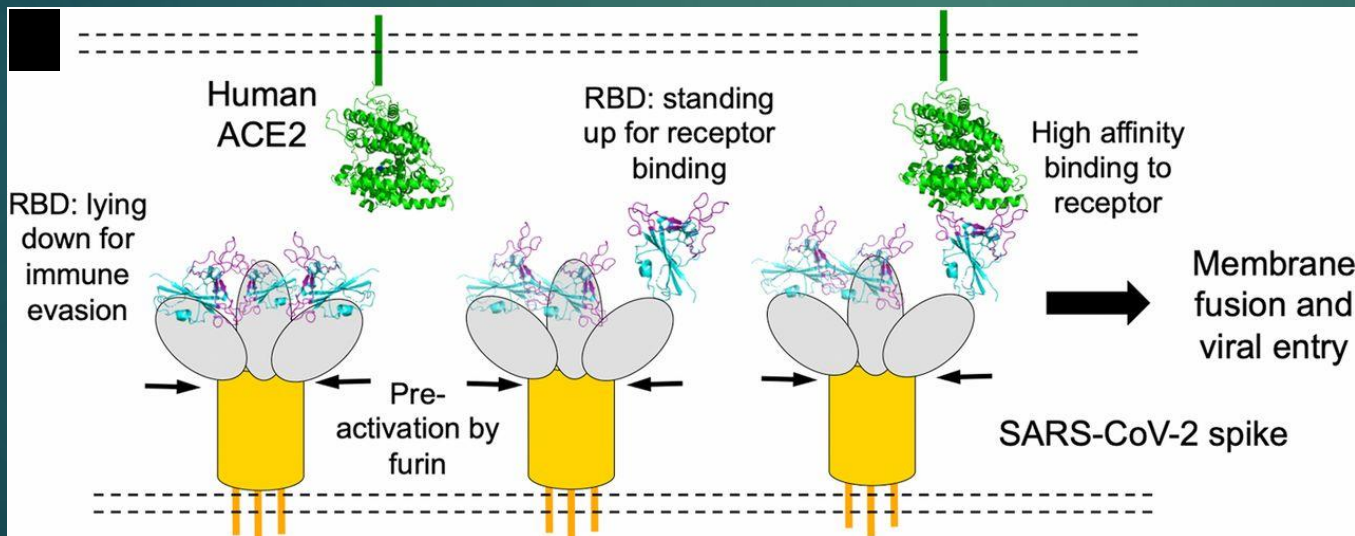
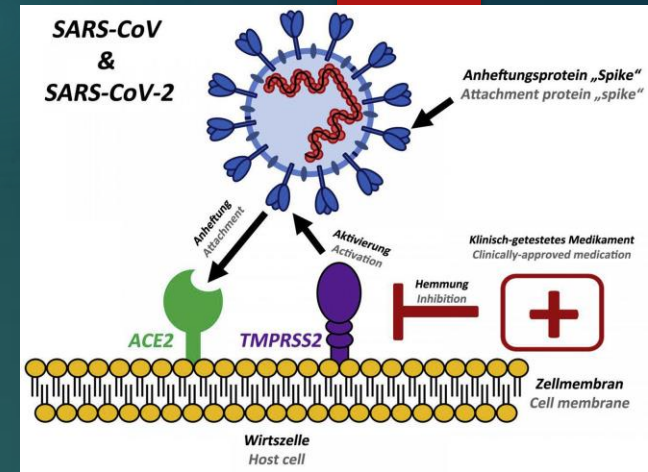
Non-structural proteins are involved in control and regulation of cellular processes.

Functions of the membrane proteins

- ▶ **M**: has a structural role in forming the virus membrane, but also can form dimers to form an ion-channel called a viroporin channel in the host cell membrane; this permits entry of Ca^+ which alters many cell processes, aids viral capsule formation and leads to apoptosis (cell death and release of viral particles).
- ▶ Expressed on the cell's membrane, M can dimerize with adjacent cells to fuse the cells and allow the virus to move to new cells without exposing itself to the immune system.

Functions of the membrane proteins

- ▶ **S** (Spike): the S protein contains the RBD (receptor binding domain) that attaches to the hACE2 enzyme found on the surface of many cells in the body. hACE2 (human Angiotensin Cleavage Enzyme) is present in arterial and venous endothelial cells and arterial smooth muscle cells in all organs studied, and abundantly present in humans in the epithelia of the lung and small intestine.



*Transmembrane serine protease 2 (TMPRSS2) is a cell surface protein primarily expressed by endothelial cells across the respiratory and digestive tracts.

The RBD is the most immunoreactive site on the spike protein. SARS-CoV2 protects that site from the immune system by keeping it hidden until the spike encounters a cell membrane which expresses proteases such as furin or TMPRSS2*, which cleaves a site near the RBD and allows it to stand up and bind the hACE2 receptor.

A Quick Aside: What is ACE2?

- ▶ ACE2 stands for **A**ngiotensin **C**onverting **E**nzyme-**2**.
- ▶ Its function is to catalyse the splitting of Angiotensin into Angiotensin(1-7)
- ▶ It is part of the kidney's blood pressure regulating function:
 - ▶ **Angiotensinogen** is an inactive molecule produced in the liver and found in the blood.
 - ▶ Low blood pressure causes the kidney to secrete **renin**
 - ▶ **Renin** is an enzyme that cleaves **angiotensinogen** into **angiotensin I**
 - ▶ ACE, found on lung epithelium (among other places) converts **angiotensin I** into **angiotensin II**
 - ▶ **Angiotensin II raises** blood pressure through vasoconstriction (anyone take “ACE-inhibitors”?)
- ▶ ACE2 acts counter to this by splitting **angiotensin II** into **angiotensin (1-7)**, which **lowers** blood pressure.*

*note for physiology buffs: this is just one example of the body's extensive use of both accelerator *and* decelerator systems (“gas and brake” for effective and rapid regulation.

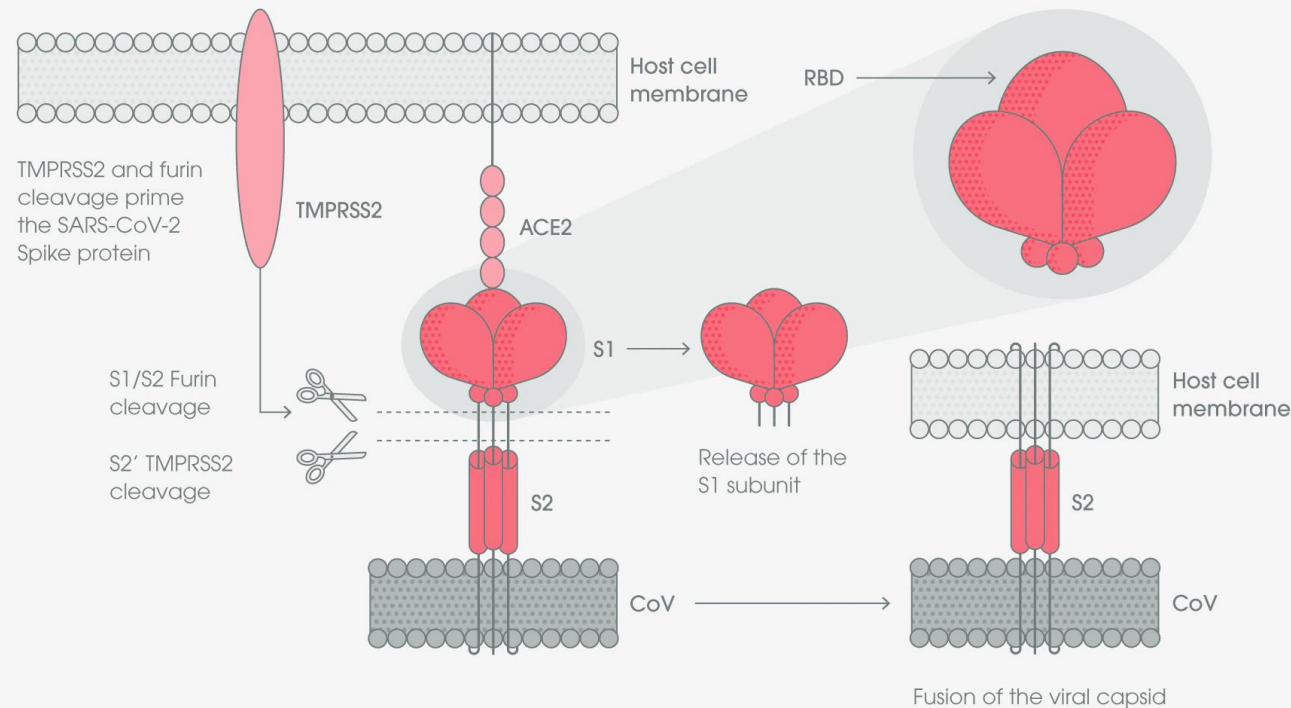
Functions of the Membrane Proteins (cont'd)

- ▶ **E:** In addition to forming the viral envelope, E plays a role in viral formation inside the affected cell. It is specific for the Golgi apparatus where many cell structures including membranes and proteins are finished and prepared for export. E causes the Golgi apparatus to perform these functions for the virus at the expense of the cell's normal processes.

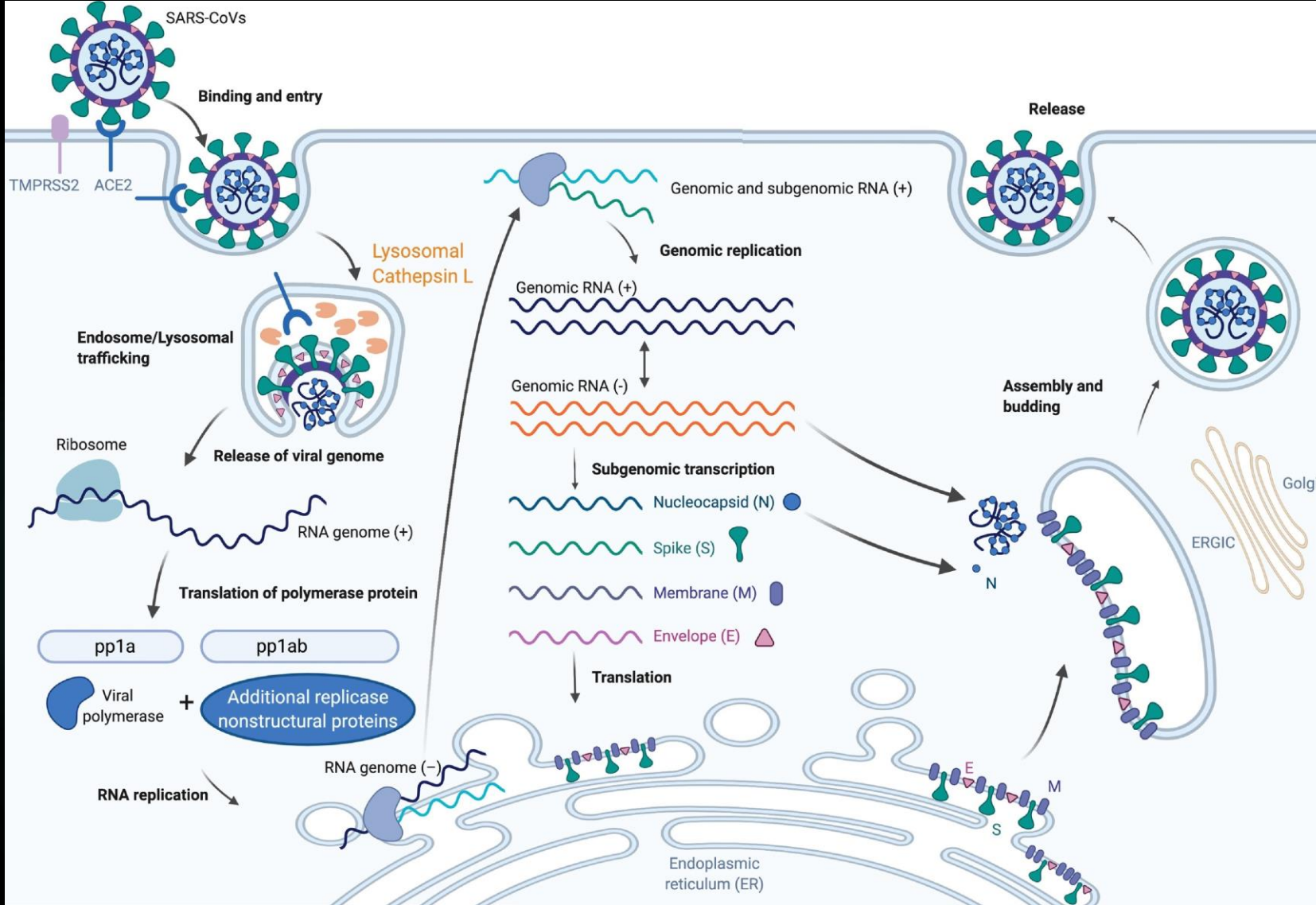
The initial attack: binding

- SARS-CoV-2 utilizes the spike glycoprotein to promote entry into the host cell. S has two functional domains

- S1 is the receptor-binding domain (RBD)
- S2 mediates the fusion of the viral and host cell membranes.



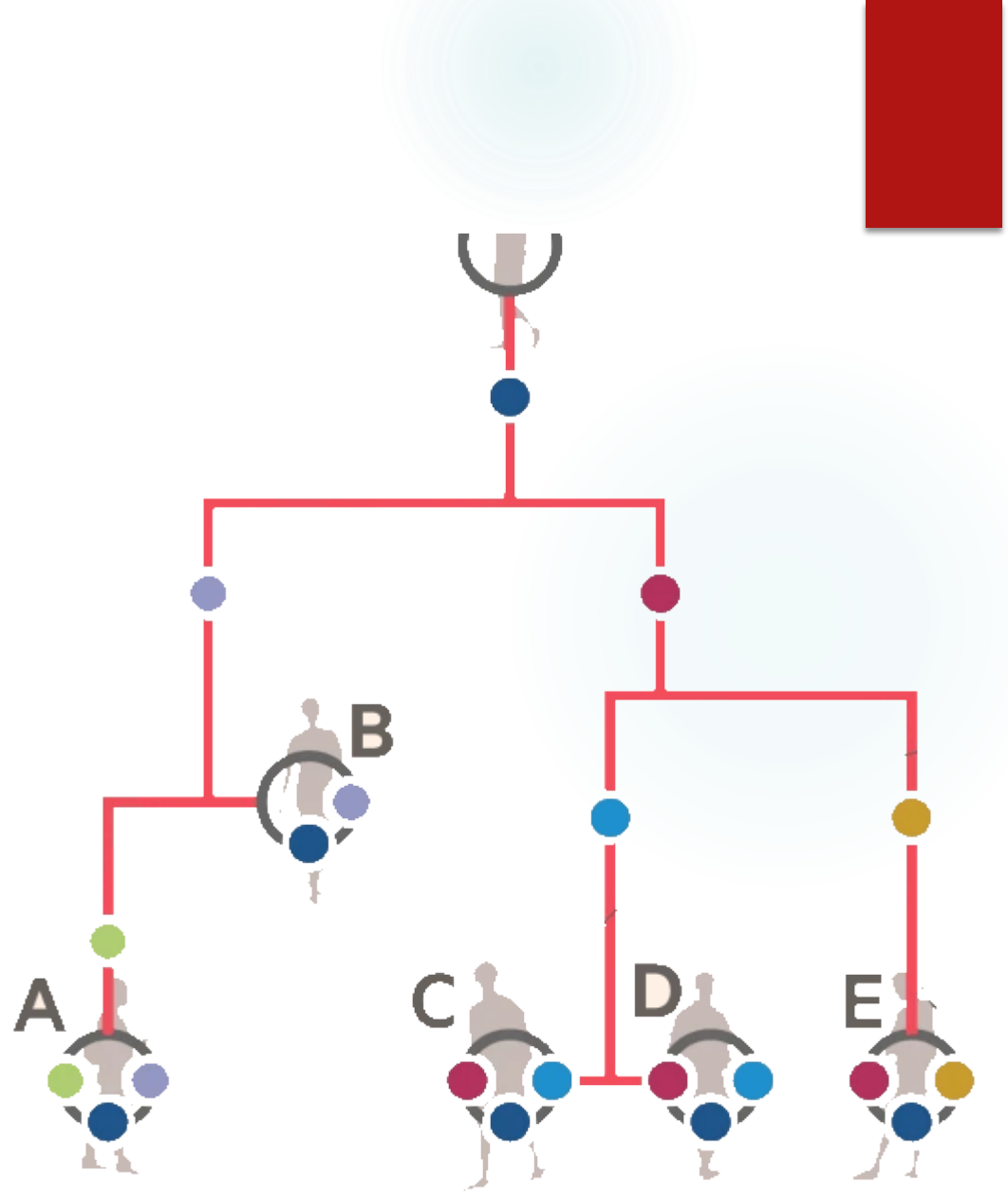
- S protein binds to the ACE2 receptor on the host cell through the S1 receptor binding domain. The S1 domain is then shed from the viral surface, allowing the S2 domain to fuse to the host cell membrane.
- This process is dependent upon activation of the S protein, by cleavage at two sites by cell surface proteolytic enzymes.



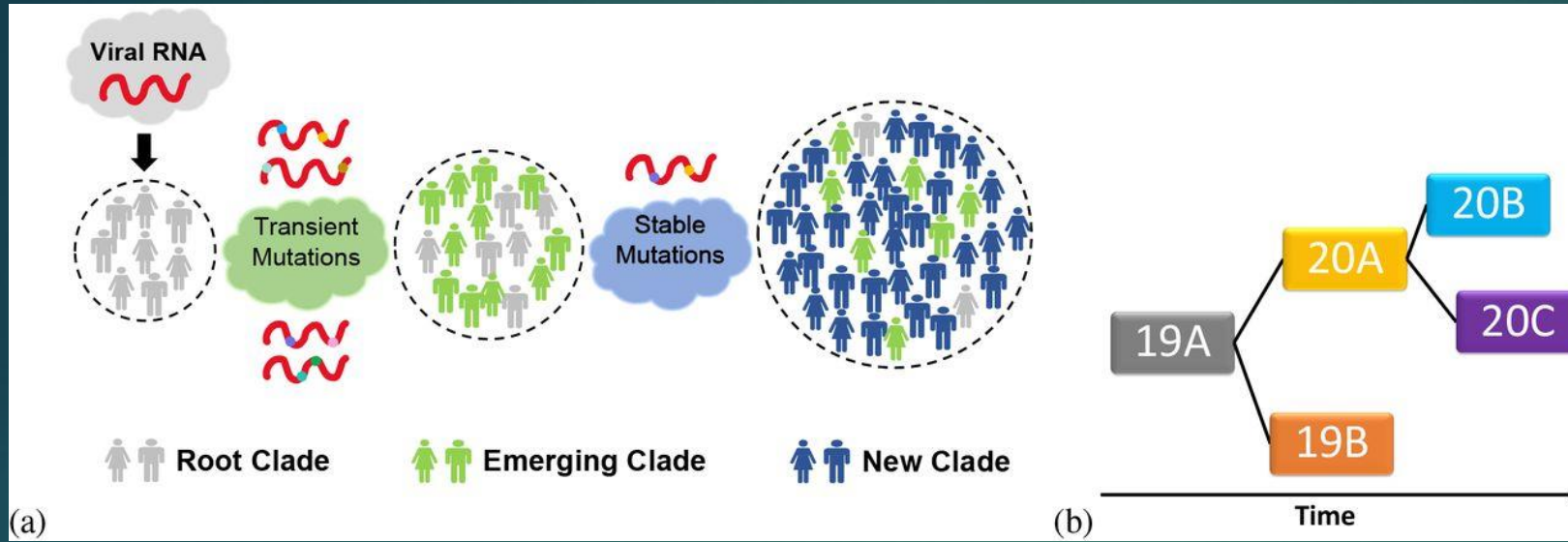
Life Cycle of the SARS-CoV2 coronavirus

Tracing mutations

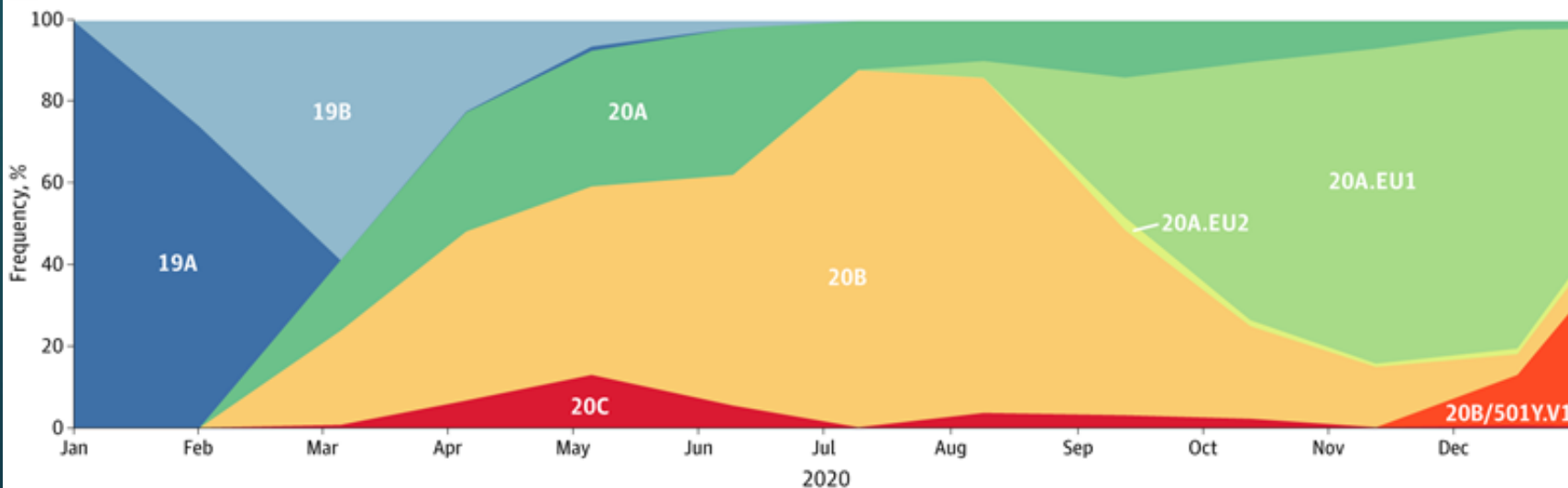
- Coronaviruses, with their extremely large genomes, have evolved a proofreading device (nsp14) which dramatically lowers mutation rate compared to other RNA viruses.
- Nonetheless they do mutate, and we can use those mutations to trace the travels of the virus.
- (nsp14 also aids in recombination with other strains of virus if two or more are infecting the same host.)



Evolution of SARS-CoV-2



B Frequencies of circulating lineages of SARS-CoV-2 over time



Evolution of SARS-CoV-2

- ▶ Why does the virus evolve so quickly?
 - ▶ Mutation Rate = # mutations/#reproduction events
 - ▶ Put another way, # mutations = mutation rate X reproduction events
 - ▶ Reproduction time for SARS-CoV2 is estimated at around 5 days*.
 - ▶ The viral load (# live viruses) an infected patient carries has been shown to be around 7×10^5 per mL in throat or sputum samples**.
 - ▶ That's seven hundred thousand viruses *per milliliter* (about 0.2 tsp), all reproducing every 5 days.
 - ▶ And that doesn't count the virus particles inside infected cells.
 - ▶ At about 500 mL fluid, that's more than three million reproductions every five days. And that's only in one patient!!

*Tapiwa et al, Eurosurveill **25** (17), March, 2020

Pan et al., Viral load of SARS-CoV-2 in Clinical Samples. The Lancet **20, 411, April 2020

Significant SARS-CoV2 mutations*

- ▶ D614G: Glycine replaces Aspartic Acid at location 614 (part of the Spike protein).
 - ▶ Outcompetes and outgrows the ancestral strain by ~10X.
 - ▶ Arose in Europe around February; 75% of infections by March.
 - ▶ Allows the cap on the spike protein to stay open more
 - ▶ This enhances attachment to ACE2 (but also opens it more to antibodies)
 - ▶ The new strain is therefore more infectious than the original (Wuhan) strain; it also replicates faster (31% faster)
 - ▶ Infected individuals show a higher viral load in the upper respiratory tract, leading to faster asymptomatic spread.
 - ▶ Today it is the dominant strain, found all over the world

*not including the original mutation(s) that allowed the virus to infect humans.

Significant SARS-CoV2 mutations

- ▶ B.1.1.7: (AKA Clade 20B or 501Y.V1) This variant, with multiple mutations, arose in South England in early December
- ▶ Mutations in the spike protein include deletions 69-70, deletion 144, N501Y, A570D, D614G, P618H, T716I, S982A, and D118H
- ▶ It is significantly more transmissible (>70%) than other variants
- ▶ It can elevate the reproductive number (R) by as much as 0.4
- ▶ In terms of severity of Covid-19, it is about the same.
- ▶ It is now found in almost every country, including the U.S.A.
 - ▶ In fact, it is here in Dutchess County.

Significant SARS-CoV2 mutations

- ▶ 501Y.V2 variant (Clade 20C)
 - ▶ First reported 18 December 2020
 - ▶ Found in the Nelson Mandela Bay area of the Eastern Cape province of South Africa
 - ▶ Mutations include N501Y, K417N and E484K
 - ▶ K417N and E484K are within the Receptor Binding Domain (RBD)
 - ▶ Spreads faster than earlier variants
 - ▶ May be causing a second outbreak in South Africa
 - ▶ Found in Switzerland on December 28 and in Japan and Australia on December 29
- ▶ Not the same as B.1.1.7 (the English variant): has N501Y, but not other characteristic mutations or the 69-70 deletion.