Bioinformatics

CS300

Chap 3

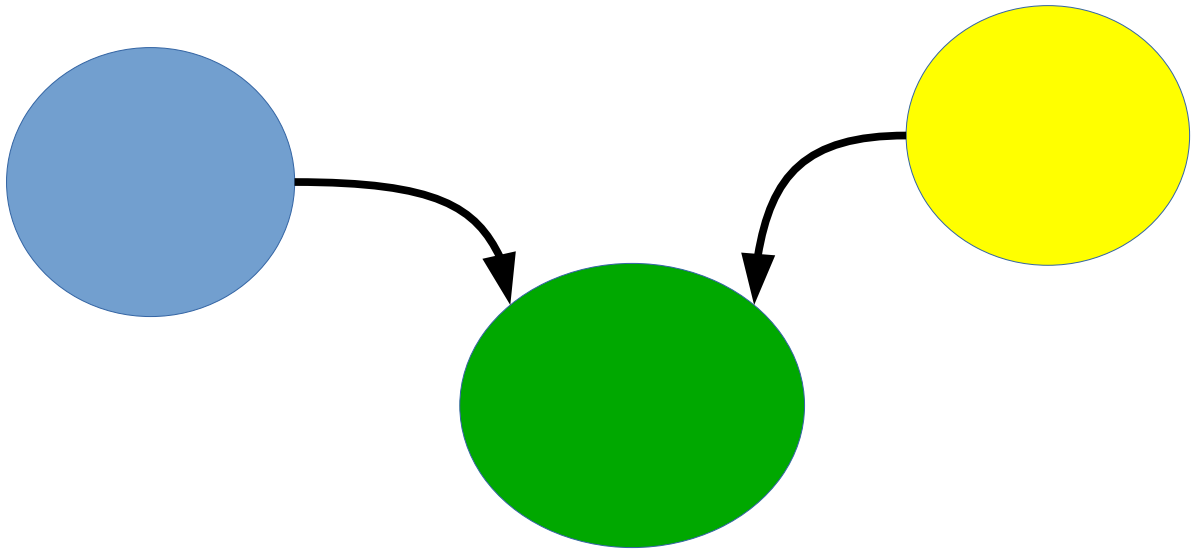
Sequence Alignment and an Influenza Outbreak

Spring 2021

Oliver BONHAM-CARTER

Descent with Modification

- ***Descent with modification*** is simply a passing trait from parent to offspring.
- One of the fundamental ideas behind Charles Darwin's theory of evolution.
- Traits are passed on to children in a process known as heredity.

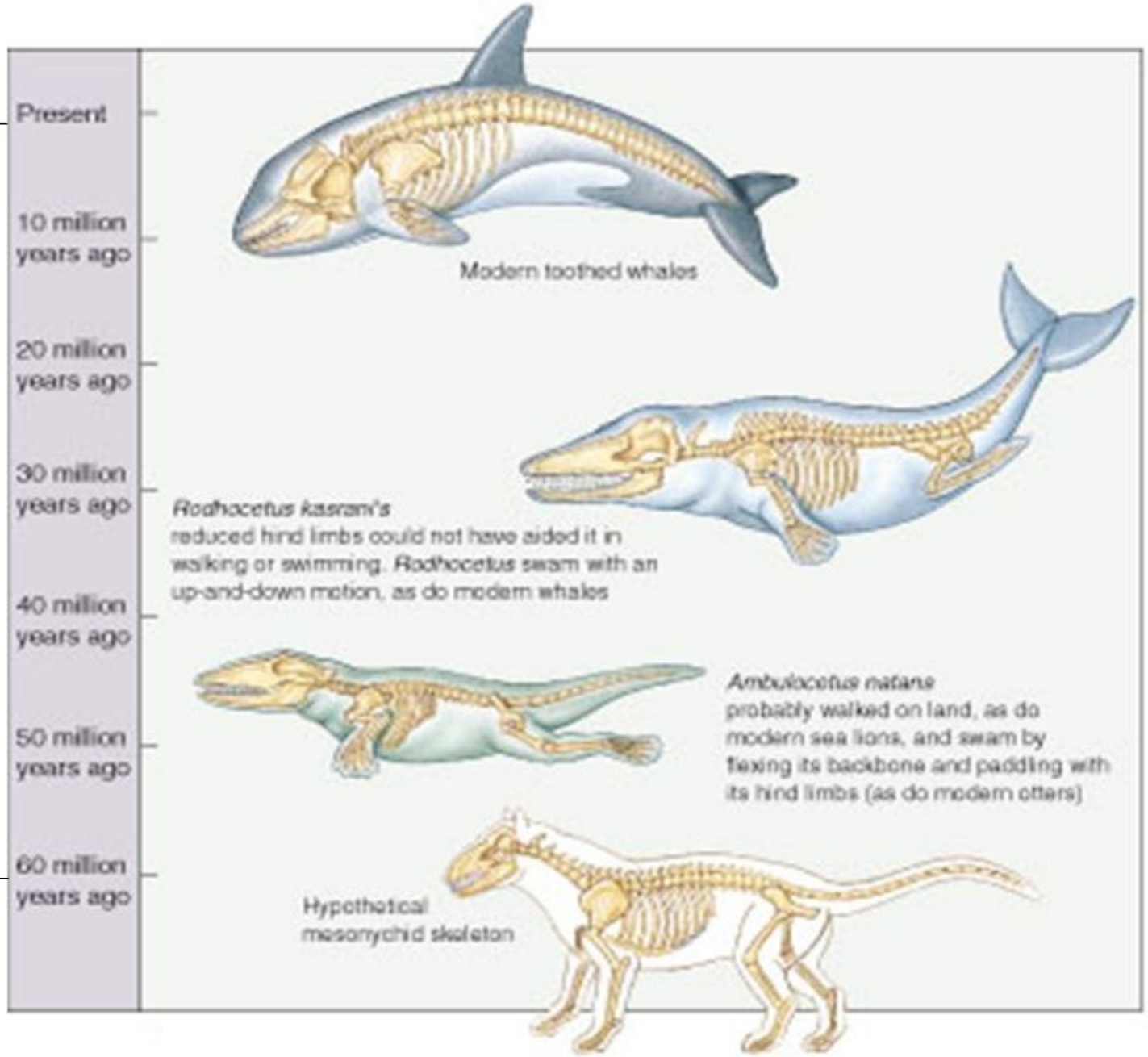


Passed Down From Genome to Genome

- DNA replication ensures a **mostly** faithful passing of the genome to progeny
- What would be the consequence of 100% accurate replication?
- Is that high similarity really desirable for a species?
- How does decent with modification happen?

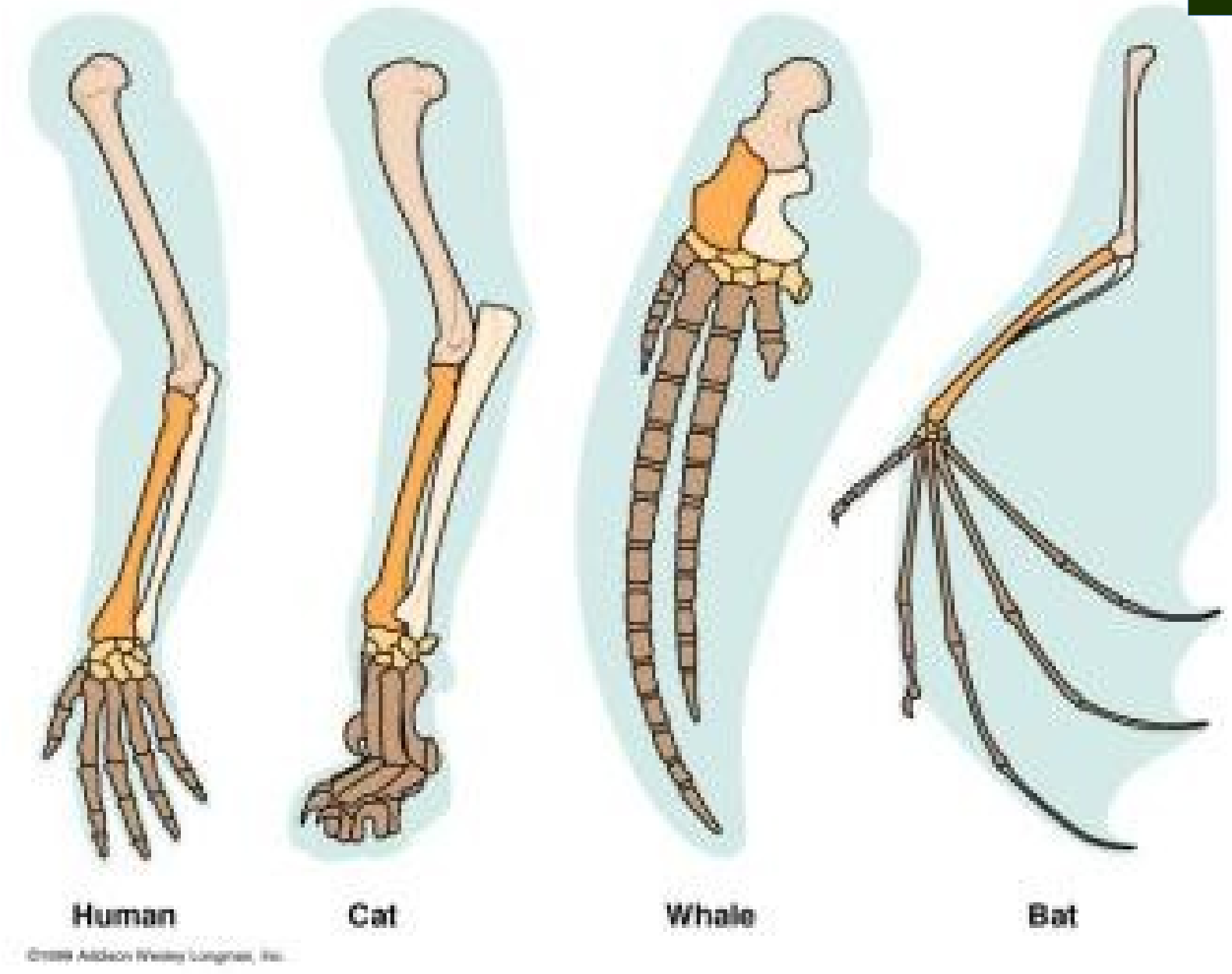
Descent With Modification

Present Day

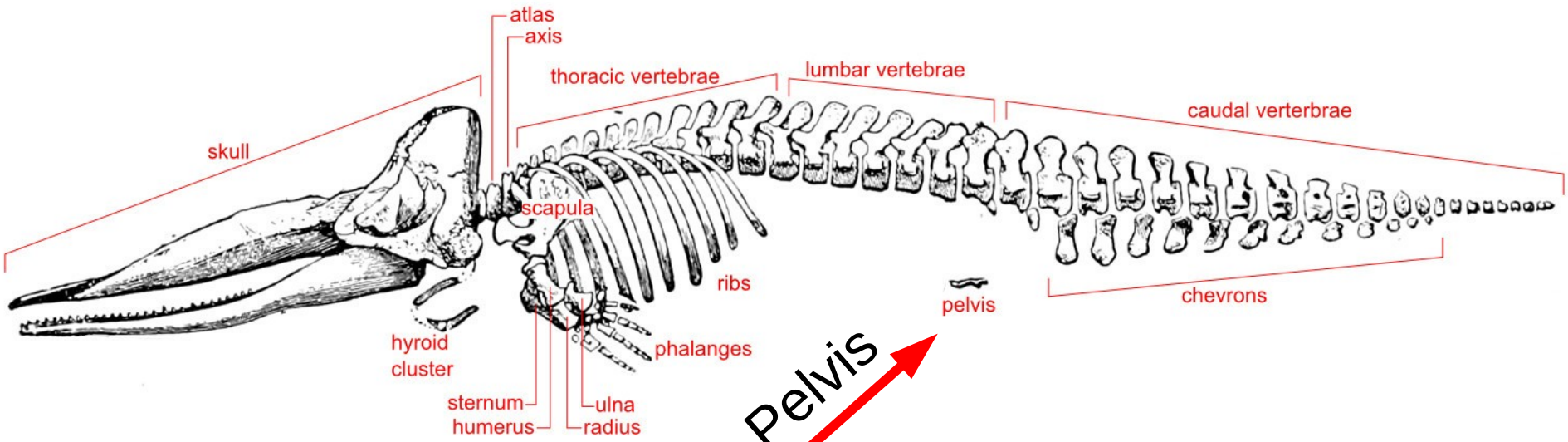


Long Ago

Same Bone, Different Day

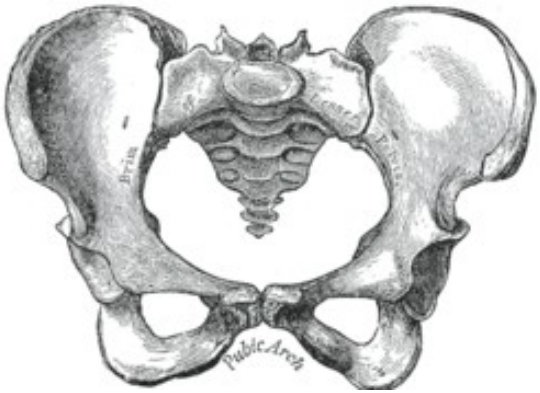


Same Bone, Different Day



Whale's Pelvis

Human's Pelvis



How Does Descent With Modification Happen?

Mutation

- A change in a DNA sequence
- Results from errors in replication or repair
- Mutation is the ultimate source of genetic variation

Domestic Dog – *Canis lupus familiaris*
All descendent from the grey wolf
All the same species (sub-species)
Breeds – variation within the species

Sequence Variations

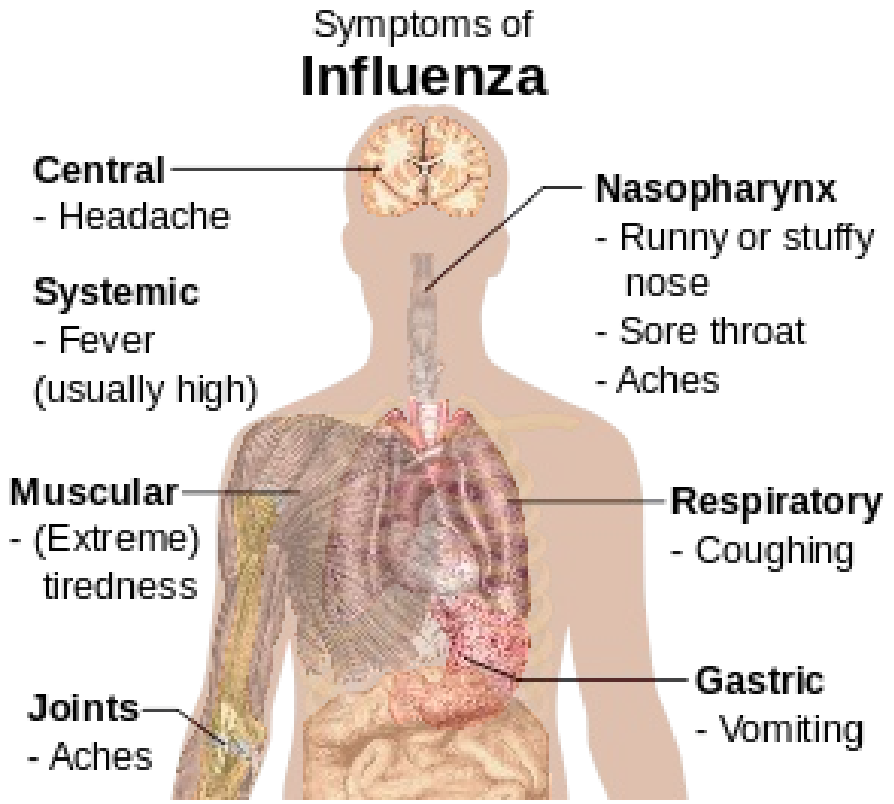
- Sequences may have diverged from a common ancestor through various types of mutations:

- substitutions (A**C**GA → A**G**GA)
- insertions (ACGA → AC**CGGA**GA)
- deletions (A**CGGA**GA → AGA)

  - You are UNIQUE and SLIGHTLY GENETICALLY DIFFERENT from each of your parents, grand parents, great grand parents …
  - Retro Viruses (Influenza, HIV, etc) are also unique and slightly genetically different from their ancesters

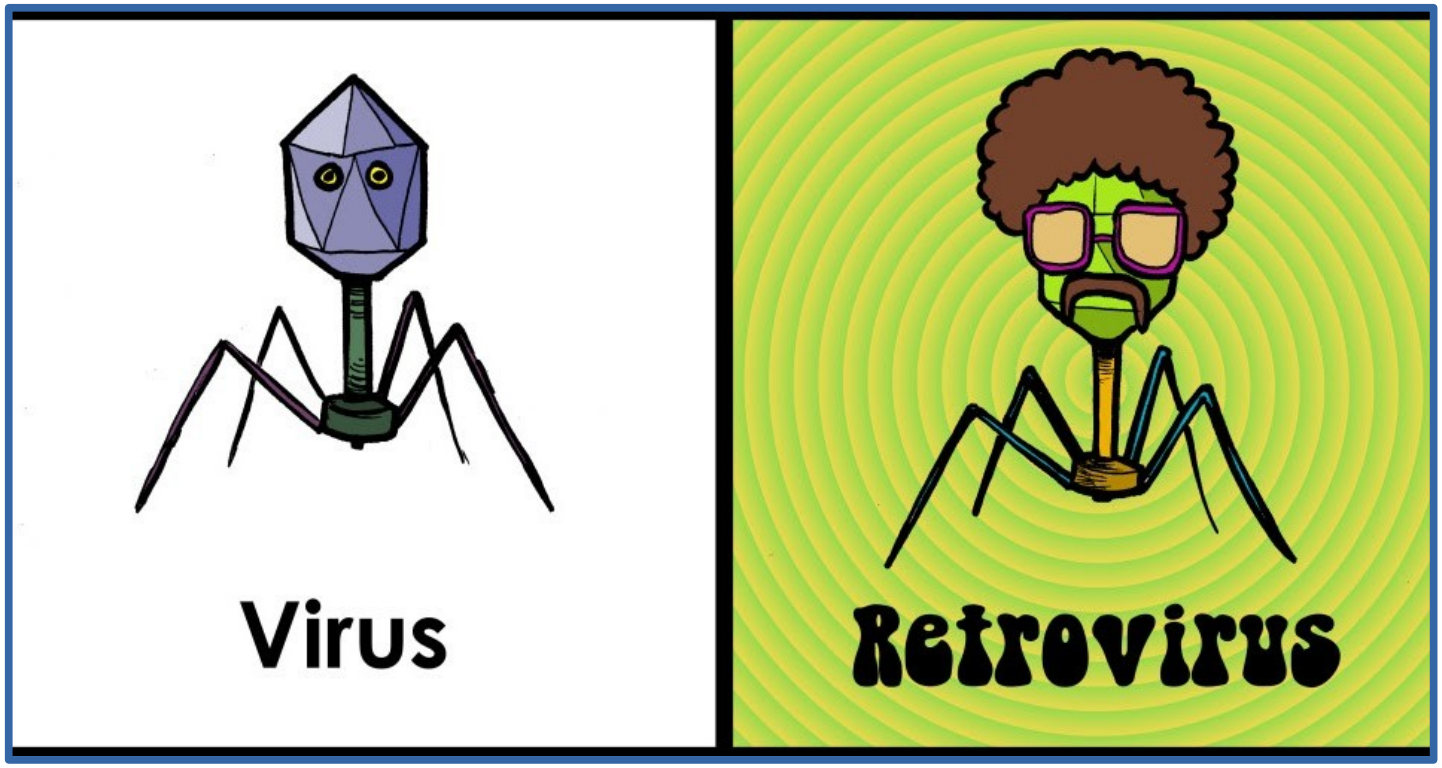
What Is Influenza?

- Flu, also known as the influenza, is a contagious disease that is caused by the flu **virus**. It attacks the respiratory tract in humans (nose, throat, and lungs). The flu is different from a cold. Flu usually comes on suddenly and may include these symptoms:
- Fever
- Headache
- Tiredness (can be extreme)
- Dry cough
- Sore throat
- Nasal congestion
- Body aches



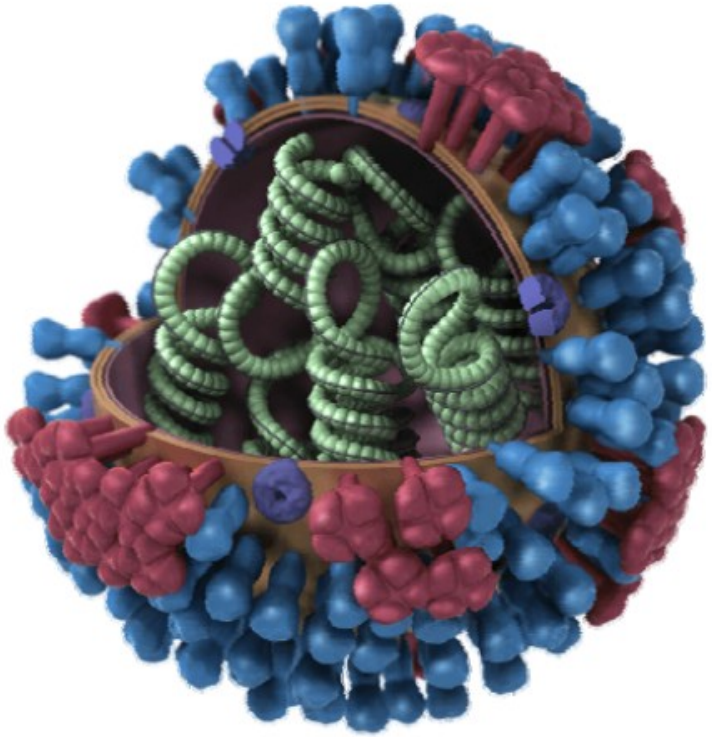
The Influenza Virus: an RNA Virus (or retro virus)

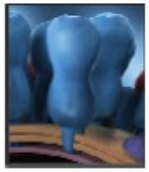
- Like all living things, influenza makes small errors—mutations—when it copies its genetic code during reproduction.
- Influenza lacks the ability to repair those errors, because it is an RNA virus; RNA, unlike DNA, lacks a self-correcting mechanism.
- As a result, influenza is not genetically stable.



H1N1 Virus

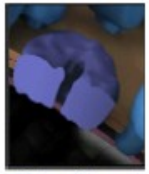
H1N1 Influenza Virus





Hemagglutinin- protein the virus uses to attach to the host cells

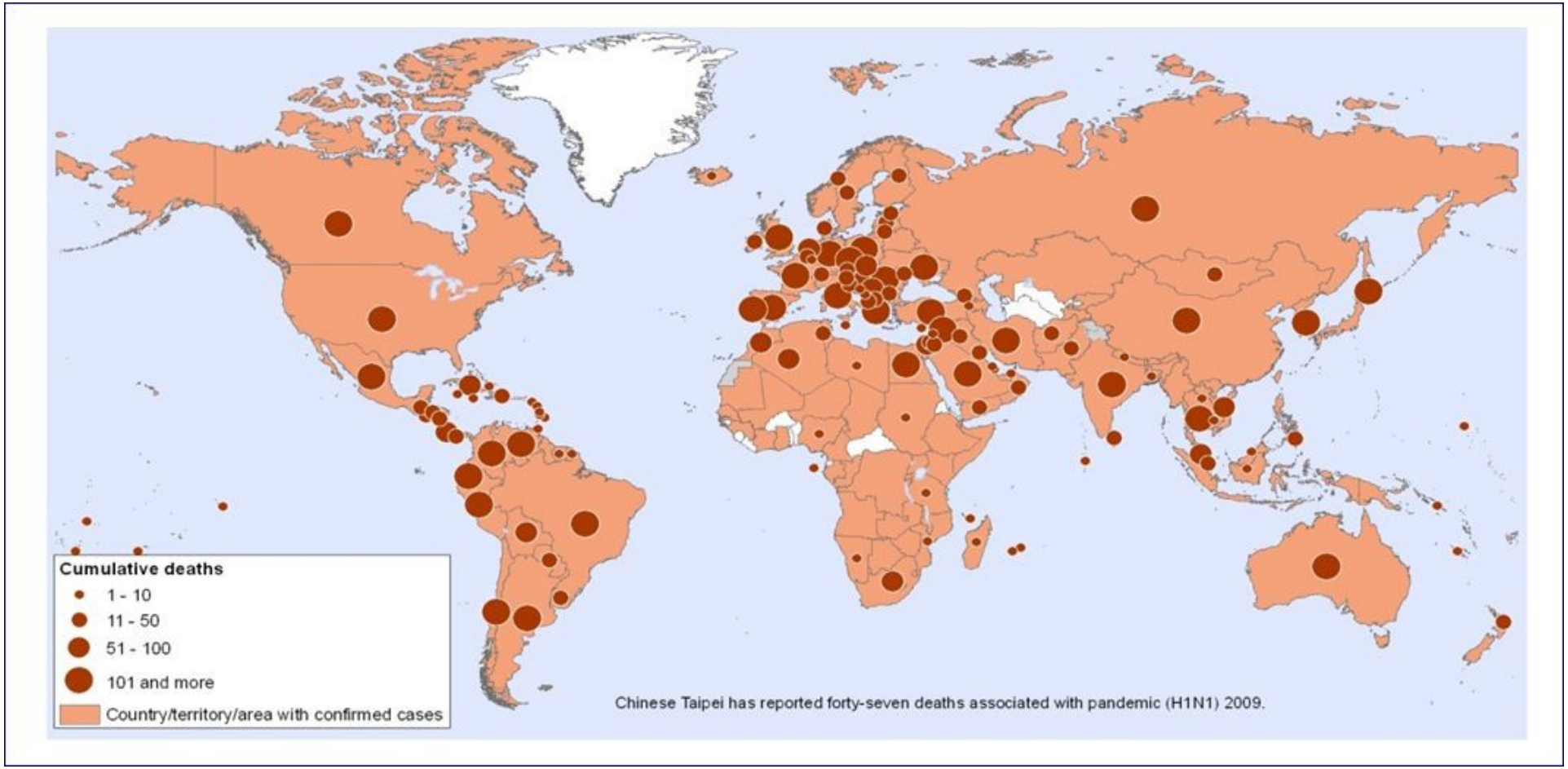
Neuraminidase- enables the virus to be released from the host cell



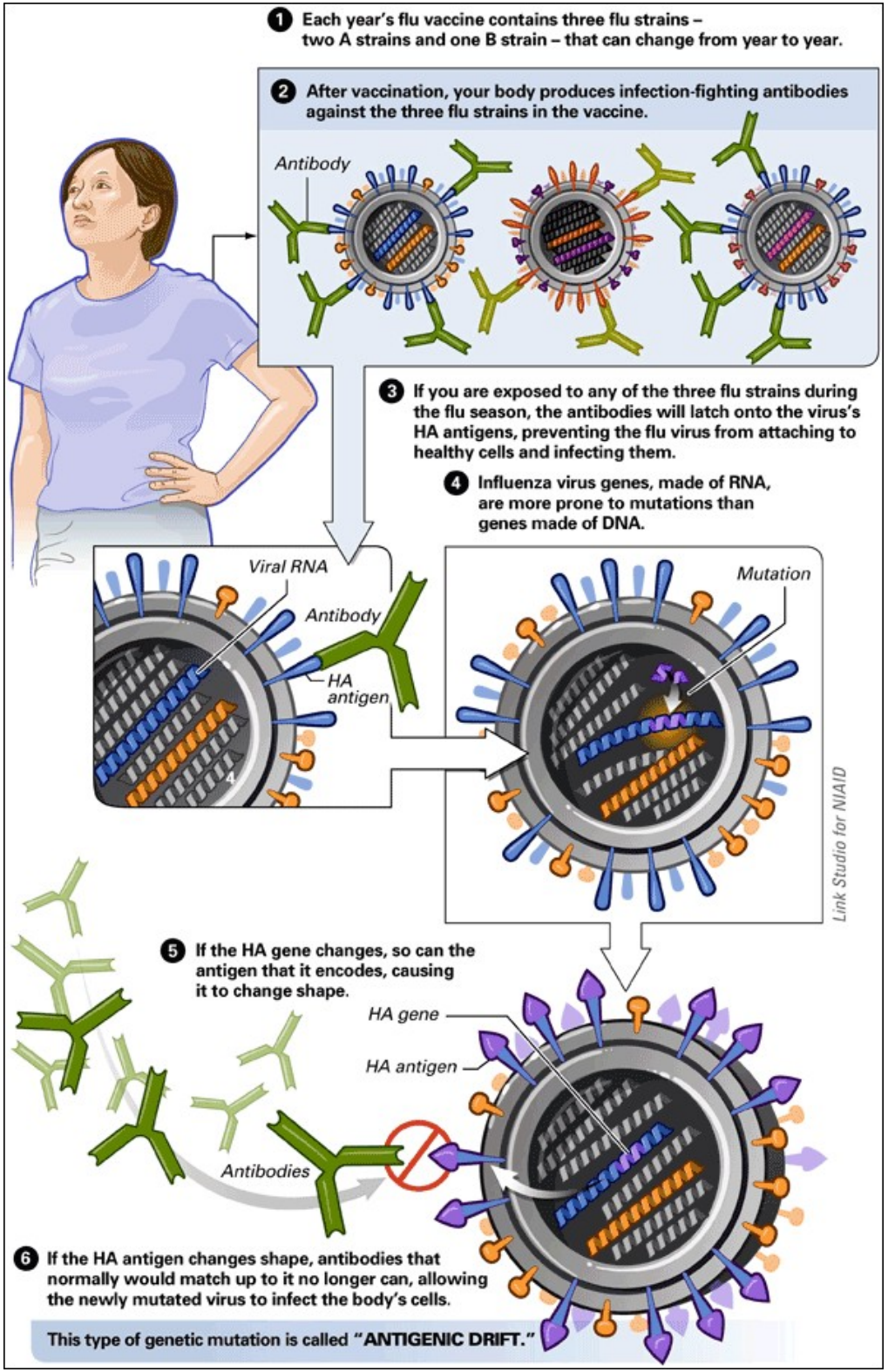
M2 Ion Channel- allows protons to move through the viral envelope and is essential for the virus replication process

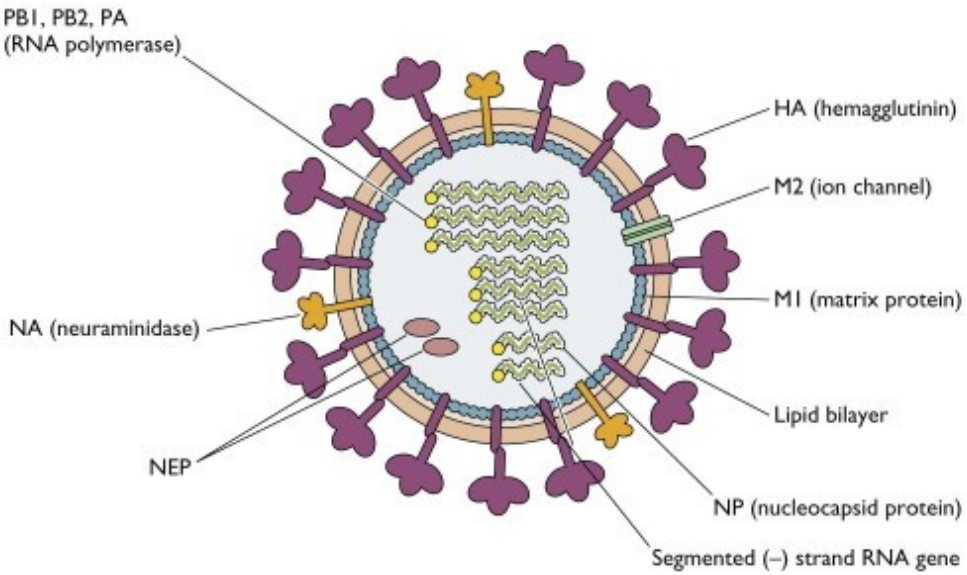
RNP- Ribonucleoprotein containing the virus RNA genome

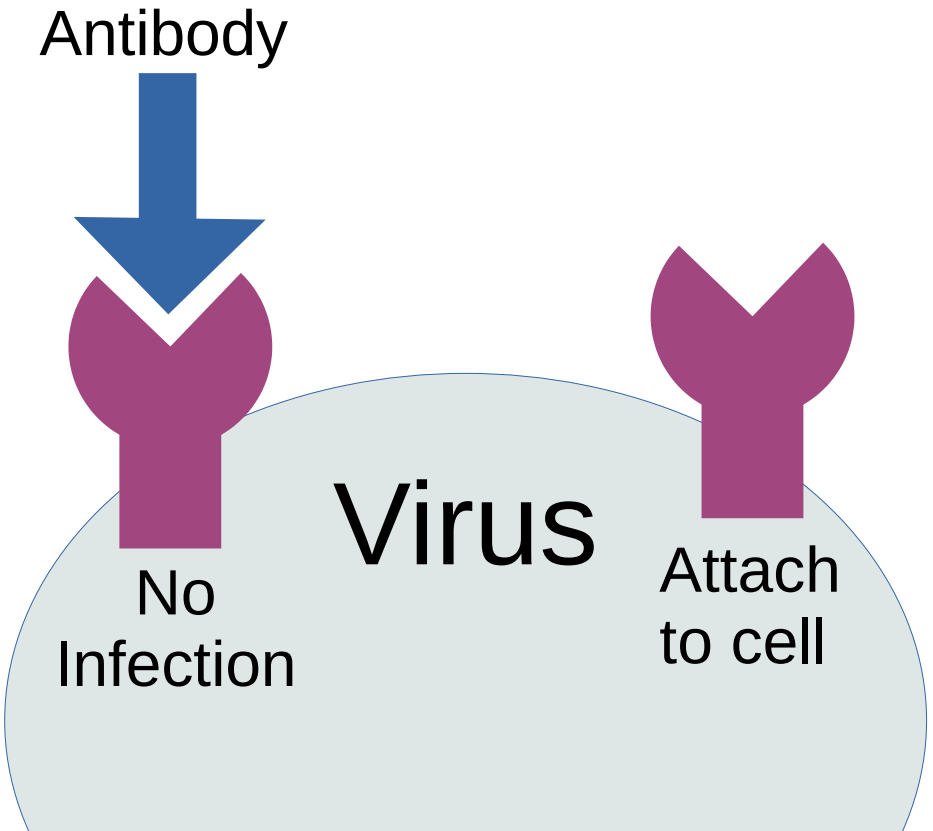
Tracking Infectious Disease
2009 H1N1 Influenza Pandemic



Vaccines


1 Each year's flu vaccine contains three flu strains – two A strains and one B strain – that can change from year to year.
2 After vaccination, your body produces infection-fighting antibodies against the three flu strains in the vaccine.
Antibody
3 If you are exposed to any of the three flu strains during the flu season, the antibodies will latch onto the virus's HA antigens, preventing the flu virus from attaching to healthy cells and infecting them.
4 Influenza virus genes, made of RNA, are more prone to mutations than genes made of DNA.
Viral RNA
Antibody
HA antigen
Mutation
Link Studio for NIAID
5 If the HA gene changes, so can the antigen that it encodes, causing it to change shape.
HA gene
HA antigen
Antibodies
6 If the HA antigen changes shape, antibodies that normally would match up to it no longer can, allowing the newly mutated virus to infect the body's cells.
This type of genetic mutation is called "ANTIGENIC DRIFT."


PB1, PB2, PA
(RNA polymerase)
HA (hemagglutinin)
M2 (ion channel)
M1 (matrix protein)
NA (neuraminidase)
Lipid bilayer
NEP
NP (nucleocapsid protein)
Segmented (–) strand RNA gene

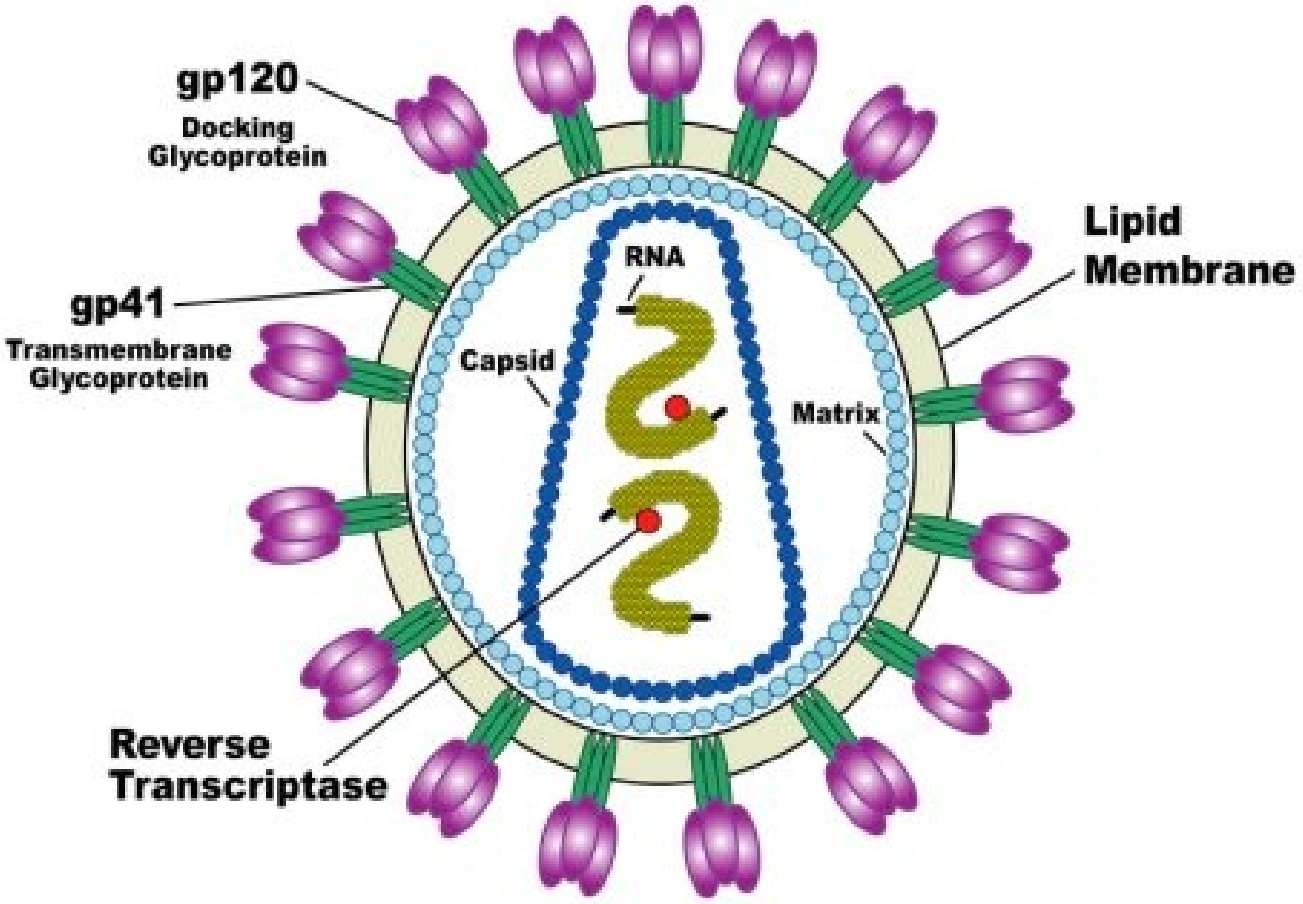

Antibody
Virus
No Infection
Attach to cell

Viral Evolution

- Viruses evolve very quickly
  - Some of the highest mutation rates known
    - Arms race with immune system
  - Viruses – mutation rate 0.0001 - 0.000001 mutations per base per generation
    - One mutation every 10,000 – 1,000,000 nucleotides
    - Influenza genome size = ~14,000 nucleotides
    - Entire genome is coding regions (genes)
  - Humans – 0.00000001 mutations per base per generation
    - One mutation every 100,000,000 nucleotides
    - Human genome size 3 billion nucleotides
    - Only 1.5% of genome is coding regions (genes)

Viruses Using DNA vs RNA

- DNA viruses are mostly double-stranded while RNA viruses are single-stranded.
- RNA mutation rate is higher than DNA mutation rate.
- Mutation causes major changes in virus genetic code
- Did mutation improve survival?



Viruses Using DNA vs RNA

- DNA replication takes place in the nucleus while RNA replication takes place in the cytoplasm.
- DNA viruses are stable while RNA viruses are unstable.
- In DNA viruses, viral genetic code is injected in the host DNA for duplication and decoding.
- RNA viruses skip DNA for duplication and decoding.

Reverse Transcription

- Retroviruses do not kill the host cell initially because they can insert their genome into the host genome.
- This process is called reverse transcription and is done by the viral protein reverse transcriptase. In the case of HIV, viral protein *integrase* then inserts the HIV DNA into host DNA.
- High mutation rate of genetic material from one organism to another.
- These mutations can be used to track virus spread.
- Sequence Alignment to compare sequences!

Molecular Epidemiology of HIV Transmissions in a Dental Practice

How related is the DNA of each virus infection?

HIV-positive persons included in study of HIV isolates from a Florida dental practice

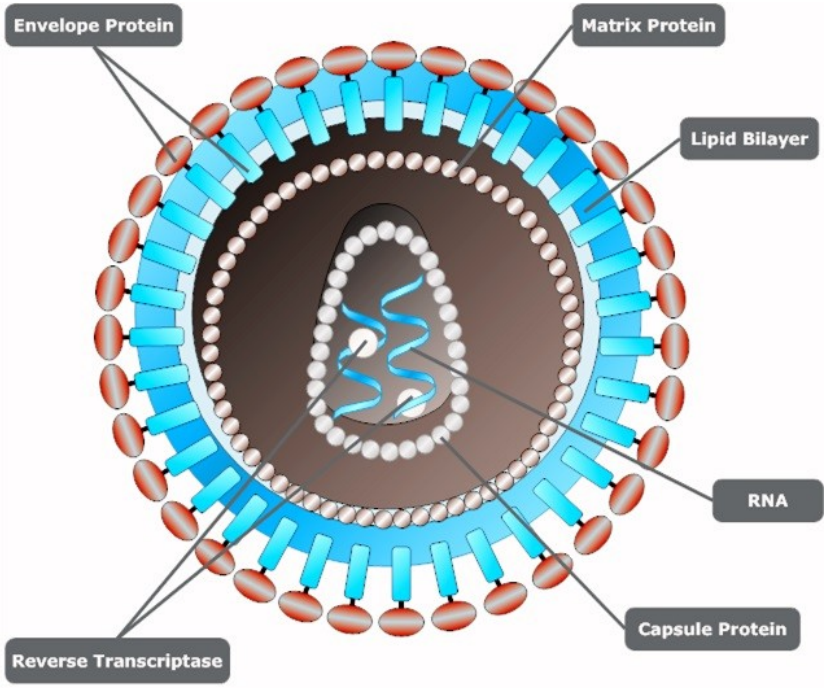
Person	Sex	Known risk factors	Average differences in DNA sequences (%) From HIV from dentist	Average differences in DNA sequences (%) From HIV from controls
Dentist	M	Yes		11.0
Patient A	F	No	3.4	10.9
Patient B	F	No	4.4	11.2
Patient C	M	No	3.4	11.1
Patient E	F	No	3.4	10.8
Patient G	M	No	4.9	11.8
Patient D	M	Yes	13.6	13.1
Patient F	**M**	**Yes**	**10.7**	**11.9**

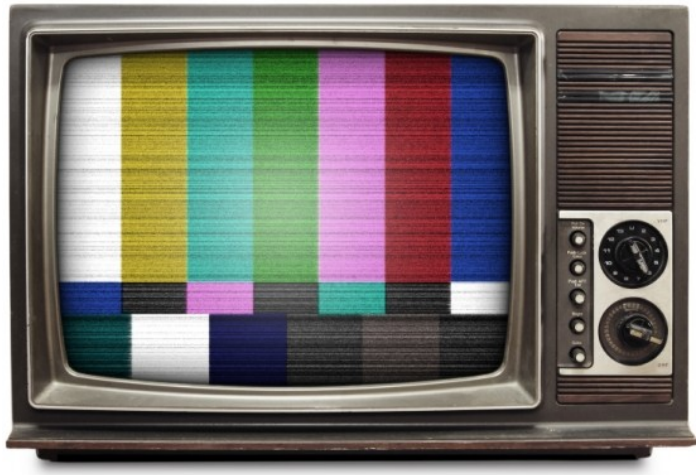
Source: After C. Ou et al., *Science* 256(1992):1165–1171, Table 1.

Similar DNA implies closer relationship between organisms

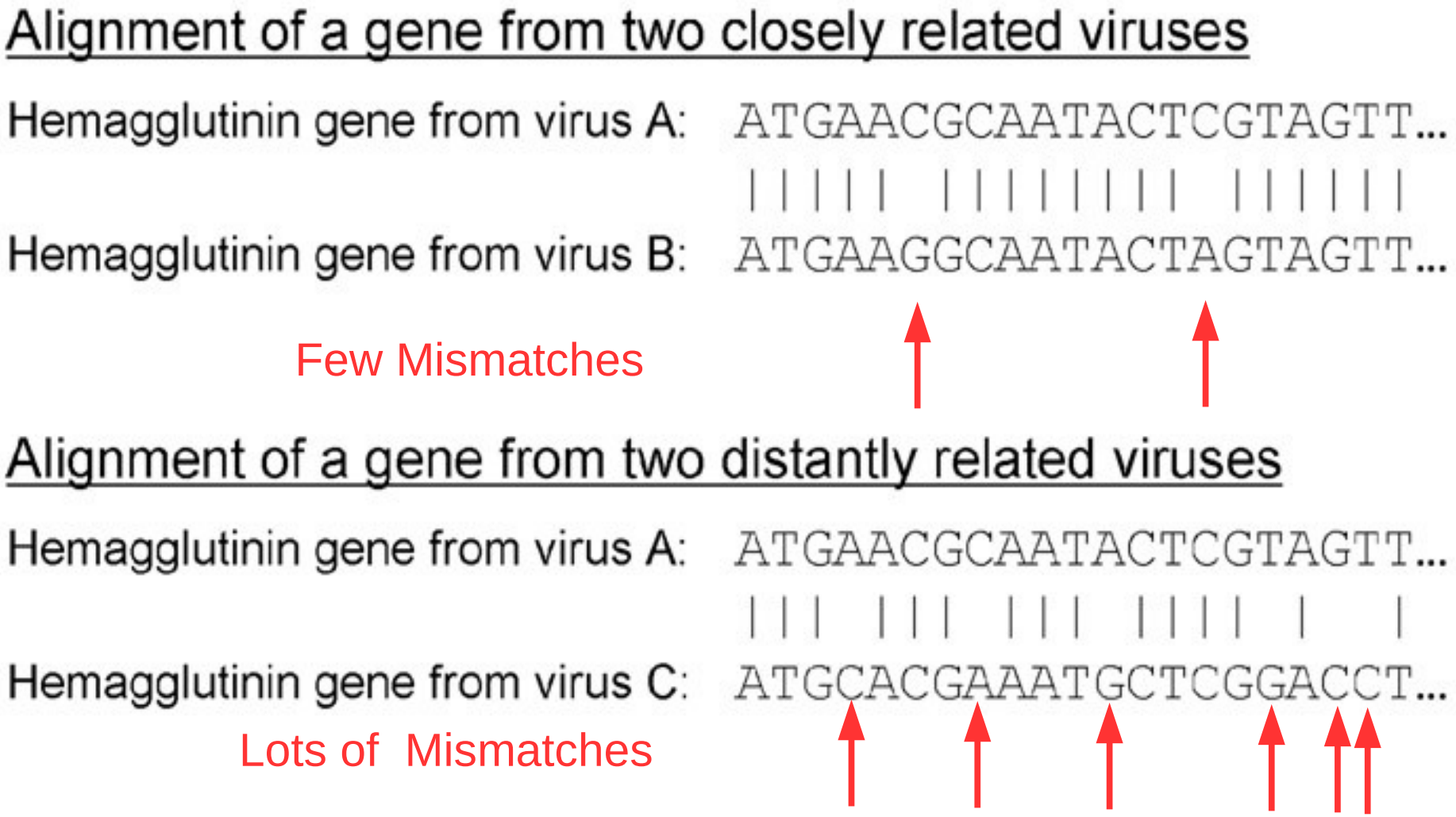
Supporting Videos

- Retrovirus reverse transcription
  - https://www.youtube.com/watch?v=eS1GODinO8w
- Coronavirus life cycle
  - https://www.youtube.com/watch?v=5DGwOJXSxqg



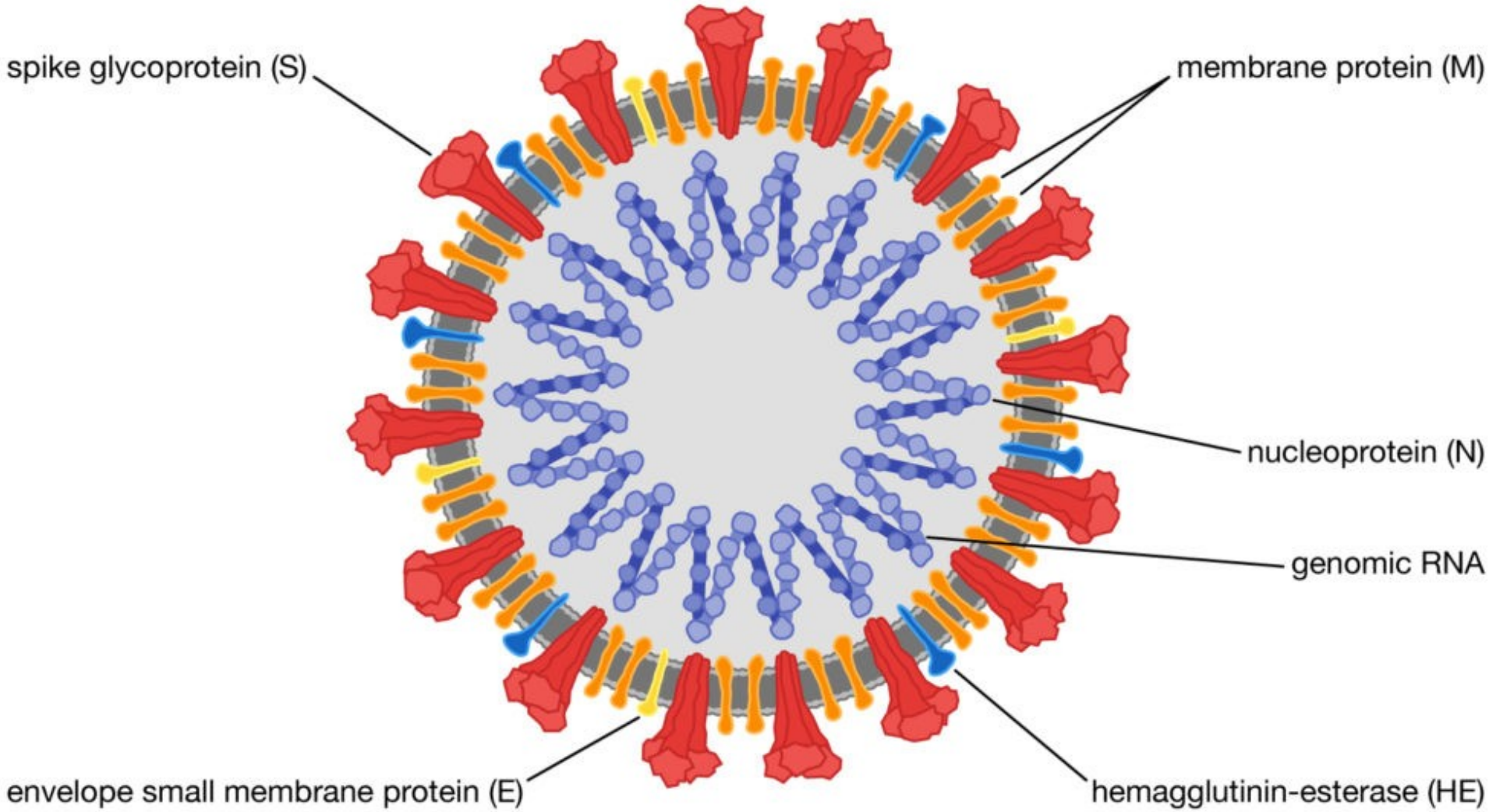


Pairwise Alignment
Similarity and Relatedness

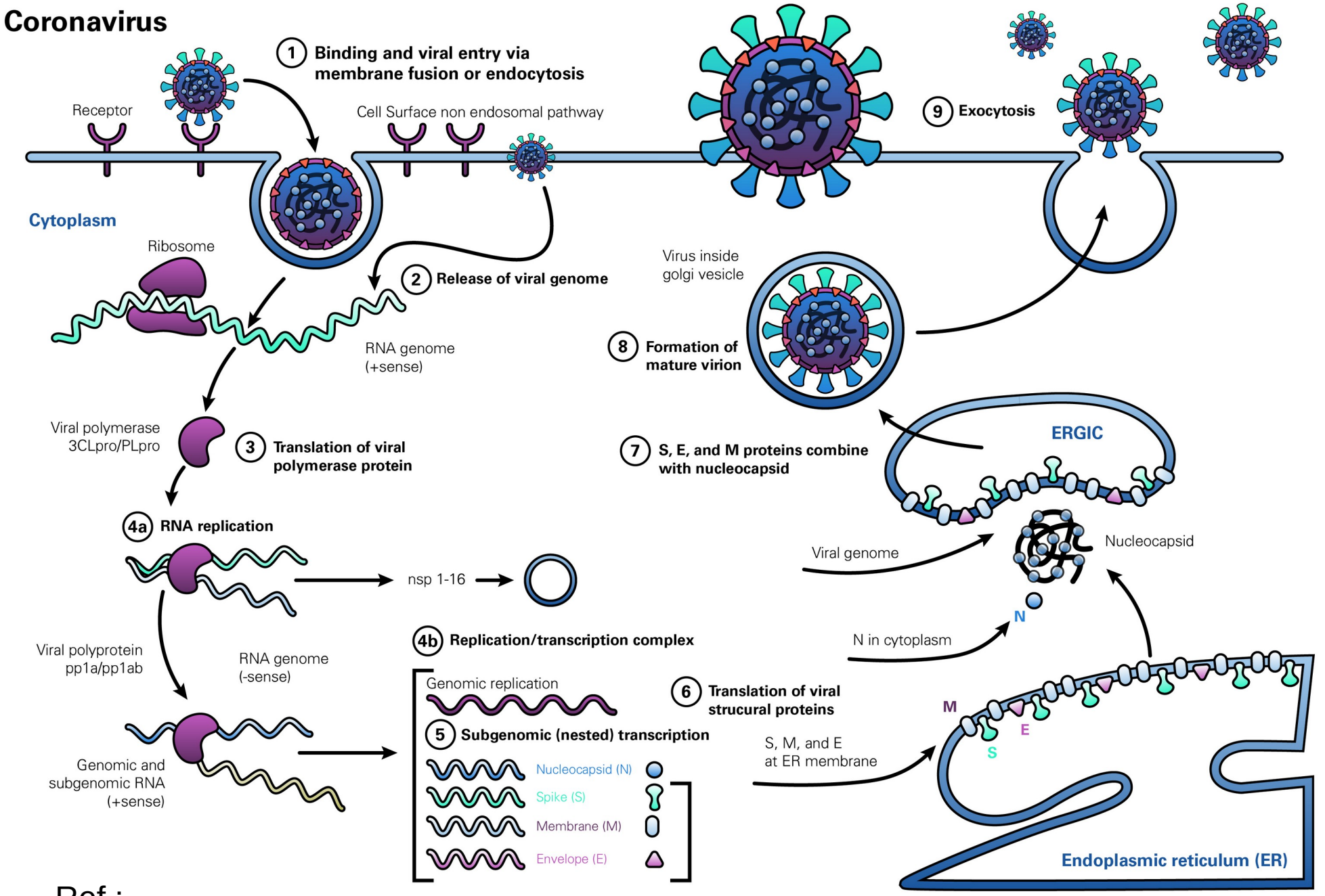


Coronaviruses

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

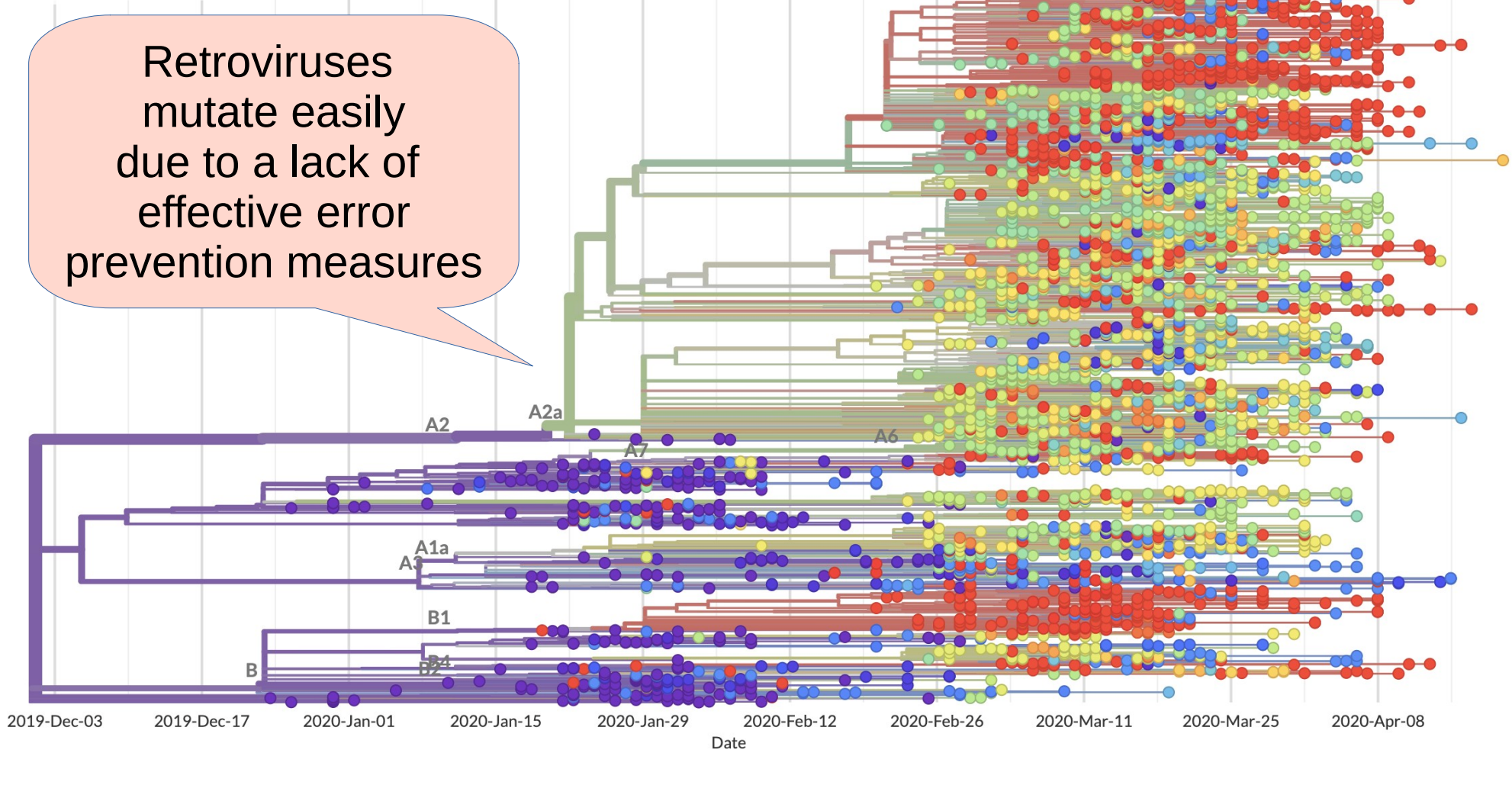


Life Cycle of Coronaviruses



Ref :
https://www.lsbio.com/media/whitepapers/sars-cov-2-and-covid-19-pathogenesis-a-review

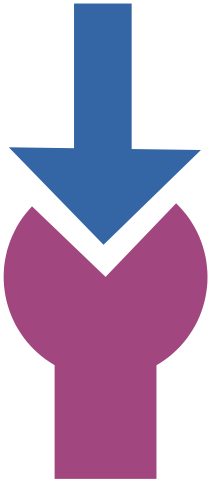
Variants of Coronavirus



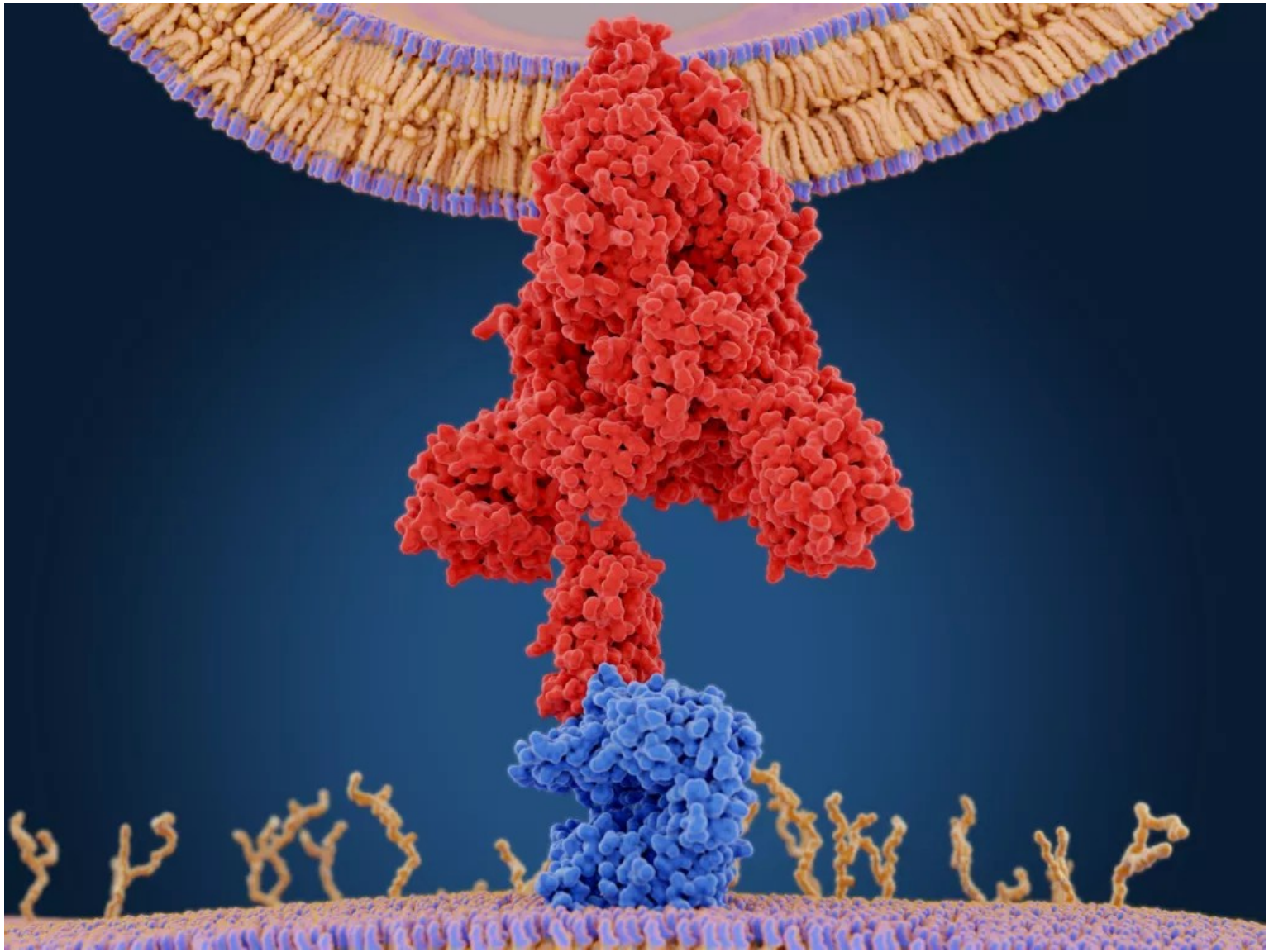
Retroviruses mutate easily by virtue of the lack of copying protections like

Spike Proteins Also Mutate...

Antibody



No Infection



Ref:
https://www.livescience.com/why-coronavirus-attaches-stronger-human-cells.htm

Midterm Reflection

- Please complete the below survey at this time.
- Your input helps to improve the quality of teaching in this course.
- Midterm reflection for CS300 spring 2021
  - https://forms.gle/D29MmgVLs1rcd2TJ8