The genetics of Cancer

Heritable mutations occur in the **germline**, which develops largely independently of the rest of the body (the soma), which, from the standpoint of heritable variation, is just a vessel[容器] for transmitting the germline

Cancers arise when **critical genes** are mutated, causing **unregulated proliferation of cells**.
These rapidly dividing cells pile up on top of each other to form a tumor.
When cells detach from the tumor and invade surrounding tissues, the tumor is **malignant** and may form secondary tumors at other locations in a process called **metastasis**[转移].
A tumor whose cells do not invade surrounding tissues is **benign**.

Tumor: a condition where there is abnormal cellular growth thus forming a **lesion** or in most cases, a **lump[肿块]** in some part of your body.
Benign tumor – grows in confined area
Malignant tumor – capable of invading surrounding tissues (metastatic process)

Cancer: degenerative disease with a cellular condition where there is **uncontrolled growing mass of cells** capable of **invading neighboring tissues** and spreading via body fluids to other parts of the body.

Names for Site of Origin
Carcinomas – epithelial cells; cover external & internal body surfaces (90%)
Sarcomas – supporting tissue; bone, cartilage, fat, connective tissue, pancreas, liver
Lymphoma & Leukemias – blood & lymphatic tissue (leukemia reserved for cancers that reside in bloodstream not as solid tissue)

Evidence of a Genetic Basis for Cancer
The cancerous state is inherited in a clonal way (无性繁殖).
Some types of **viruses** can induce the formation of tumors in experimental animals.
Cancer can be induced by mutagens.
Certain types of white blood cell cancers are associated with **chromosomal abnormalities**.

Cancer and Genes
Oncogenes: An oncogene is a gene that has the potential to **cause cancer**. In tumor cells, they are often mutated and/or expressed at high levels. Their activity in cancer onset is linked to a gain or an elevation in their function.
Tumor: A tumor suppressor gene, or anti-oncogene, is a gene that protects a cell from one step on the path to cancer. When this gene mutates to cause a loss or reduction in its function, the cell can progress to cancer, usually in combination with other genetic changes.

Tumor-Inducing Retroviruses and Viral Oncogenes

Retroviruses have an RNA genome.

The **Rous sarcoma virus**, the first tumor-inducing virus, contains four genes

– gag encodes the capsid protein of the virus

– pol encodes the reverse transcriptase

– env encodes a viral envelope protein

– v-src encodes a protein kinase that inserts into the plasma membranes of infected cells.

– The v-src gene is an **oncogene** that is responsible for the virus's ability to induce abnormal cell growth.

Proteins Encoded by Viral Oncogenes: Growth factors similar to those encoded by cellular genes or proteins similar to growth-factor and hormone receptors; Tyrosine kinases that do not span the plasma membrane, Transcription factors homologous to cellular proteins.

Tumor Suppressor Genes

Many cancers involve the inactivation of genes whose products play important roles in regulating the cell cycle.

C-ras and c-myc: genes required for regulation cell cycle.

- increase activity and/or concentration-----oncogene----may form tumors.
- decrease activity and/or concentration-----anti-oncogene-----favors tumor formation

Detecting the frequency of new mutants by flow cytometry using antibodies specific for glycosylphosphatidylinositol-linked proteins (e.g., CD48, CD55, and CD59)

Role of inherited mutations in cancer

One is obvious: Mutations that result in an **elevated rate of somatic mutations**

the other one is linked to us being **diploid** and therefore the two alleles of an anti-oncogene need to be mutated to develop a cancer (will elaborate in a few slides)

pBRCA1 and pBRCA2 regulate DNA repair

Mutations in the tumor suppressor genes BRCA1 (Ch17) and BRCA2 (Ch13) have been implicated in hereditary breast and ovarian cancer.

Both genes encode proteins that are localized in the nucleus and have putative transcriptional activation domains.

pBRCA1 and pBRCA2 may be involved in DNA repair in human cells.

Knudson's Two-Hit Hypothesis

When tumor suppressor genes are mutated, a predisposition to develop cancer often follows a **dominant** pattern of inheritance.

The mutation is usually a loss-of-function mutation in the tumor suppressor gene.

Cancer develops only if a second mutation in somatic cells knocks out the function of the wild-type allele.

Cellular Roles of Tumor Suppressor Proteins

The proteins encoded by tumor suppressor genes are involved in:

cell division, cell differentiation, programmed cell death, DNA repair.

Cancers develop through **an accumulation of somatic (not a single) mutations** in proto-oncogenes and tumor suppressor genes.

Multiple Mutations in Cancer

Most malignant tumors cannot be attributed to mutation of a single gene.

Tumor formation, growth, and metastasis depend on the accumulation of mutations in several different genes.

The genetic pathways to cancer are diverse and complex.

Somatic mutation is the basis for the development and progression of all types of cancer.

As mutations accumulate and cells become unregulated, genetic instability increases the likelihood that the cells will develop the hallmarks of cancer.

Cell biology of cancer

Cancer cells share 2 properties that differentiate them from normal cells

1. The ability to proliferate (reproduce/divide) in the absence of normal signals, often indefinitely
2. The ability to invade environments/tissues where they don't belong

A small number of key regulatory pathways are perturbed in almost all human cancers. These pathways regulate cell proliferation, cell growth, cell survival, and the cell's response to DNA damage or stress, like alterations in cell proliferation, DNA damage response, cell growth, and cell survival.

Cancer cells have typically lost multiple regulatory pathways that enable normal cells to respond appropriately to their environment. Mutations in genes affecting key regulatory pathways can lead cells to over-proliferate or fail to respond appropriately to negative growth signals.

Gain of function mutations in proto-oncogenes and/or loss-of-function mutations in tumor suppressors can lead to cancer.

TABLE 20–2 EXAMPLES OF CANCER-CRITICAL GENES

Gene	Class	Effect of Mutation
Ras	Proto-oncogene	Activating mutations in *Ras* render the Ras protein continuously active, promoting cell proliferation (discussed in Chapter 16)
β-catenin	Proto-oncogene	Activating mutations in *β-catenin* make the β-catenin protein resistant to degradation, promoting cell proliferation (see How We Know, pp. 730–731)
p53	Tumor suppressor gene	Inactivation of both copies of *p53* allows cancer cells to continue to survive and divide, even in the presence of damaged DNA (discussed in Chapter 18)
APC	Tumor suppressor gene	Inactivation of both copies of *APC* promotes excessive proliferation of cells in the intestinal crypt (see Figure 20–52 and How We Know, pp. 730–731)
Brca1 and *Brca2*	Tumor suppressor genes	Inactivation of both copies of *Brca1* or *Brca2* allows cancer cells to continue to survive and divide in the presence of massively damaged DNA (discussed below).

v-Src was the first oncogene discovered. It is carried by a bird virus (Rous sarcoma virus). The normal, cellular version is c-Src. It is a "nonreceptor tyrosine kinase." Mammalian Src proteins are also frequently mutated in cancer. v-Src protein is missing a C-terminal region that contains an inhibitory phosphorylation site (and is thus a gain-of-function allele).

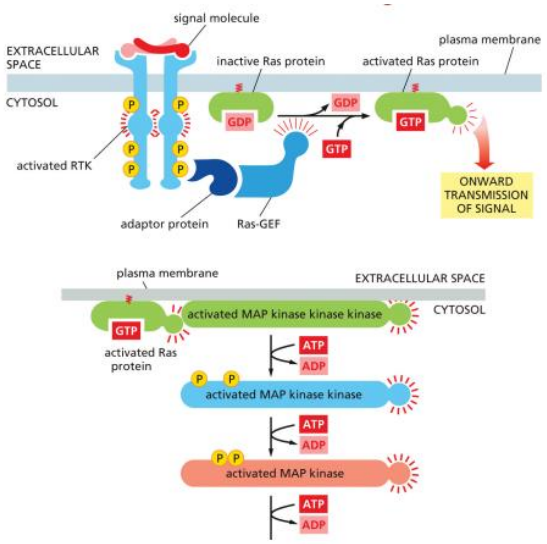
Most cancers are caused by spontaneous mutations (as described by Xavier). However, most cervical cancer, and many oral and anal cancers, are caused by HPV (human papilloma virus) infection. The HPV genes E6 and E7 are oncogenes. They inactivate p53 and Rb, respectively, in infected cells.

Gain of function mutations in proto-oncogenes and/or loss-of-function mutations in tumor suppressors can lead to cancer. Ras is the most commonly mutated gene (oncogene) in human cancers.

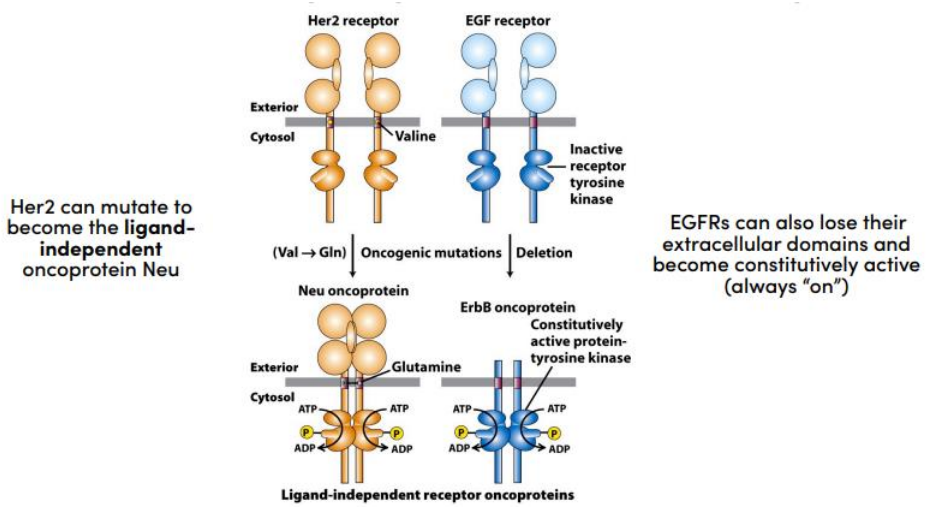
Many cellular growth factors bind to Receptor Tyrosine Kinases (RTKs), which signal through Ras GTPases. Any mutation in Ras that reduces its GTPase activity will hyperactivate its signaling.

Note: these are GAlN-of-FUNCTION mutations, even though they seem to reduce a function of Ras.

Loss-of-function mutations in Ras GAP shave the same effect, and Ras-GAPs arethus regarded as tumor suppressors. NF1 (neurofibromatosis 1) is a Ras GAP

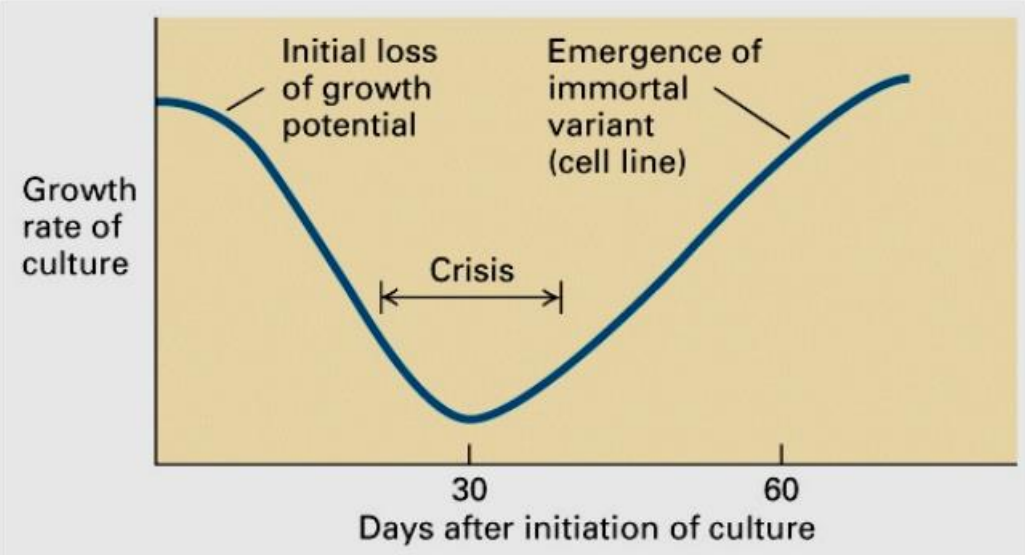


Some RTKs, particularly EGFRs (epidermal growth factor receptors) can be mutated in cancer, usually in ways that make them constitutively active.



Embryonic stem cells are “totipotent”: They give rise to all types of cells in an organism. Stem cells divide to give rise to both new stem cells (“self-renewal”) and cells that undergo terminal differentiation

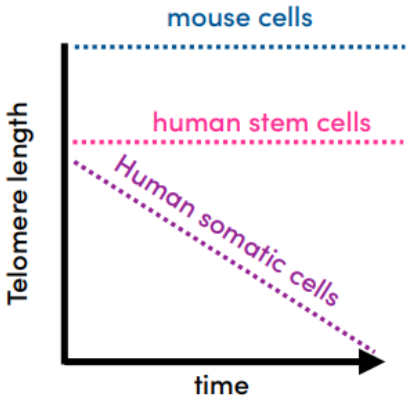
Most cells in the body or removed from the body are MORTAL: they have a finite ability to replicate and divide.



Exceptions: embryonic stem cells and cancer cells

Telomeres - the ends of eukaryotic chromosomes, have to be replicated in a special way.
TELOMERES - unique sequences at chromosome ends enable complete replication of chromosomes; also recruit SHELTERIN to protect chromosome ends from fusion/recombination
TELOMERASE - An enzyme that adds bases to the ends of chromosomes to prevent erosion of telomere sequences

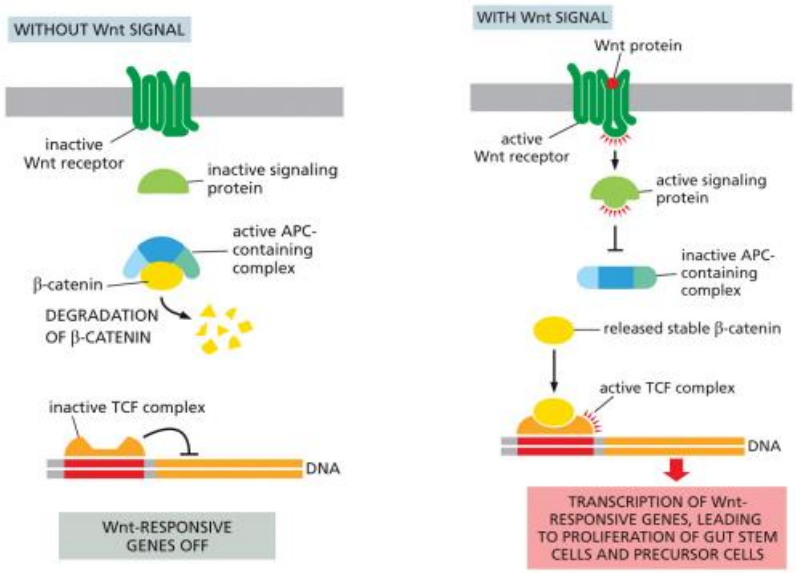
Lagging strand synthesis cannot reach all the way to the end of the chromosome, approximately 50-70 bases are lost each replication cycle.
In stem cells, telomere length is maintained by telomerase, an enzyme complex that extends them to compensate for incomplete replication. In humans, telomerase is inactivated as cells differentiate.



Cancer cells often reactivate telomerase, or acquire an ALTernate mechanism to maintain telomeres through recombination, called ALT

Mutation of TUMOR SUPPRESSORS can lead to genome instability, which in turn can promote the acquisition of multiple mutations that lead to cancer.

The APC* protein normally keeps the Wnt signaling pathway inactive when the cell is not exposed to a secreted Wnt signal protein.



Mutation of ß-catenin to a degradation-resistant form can have a similar effect to loss-of-function of APC.

Cancer genomics

1.Cancer is not just one disease
2.Genomics can help reveal the diversity of mutations underlying different cancers
3.Using cancer genomics can aid in personalizing medic

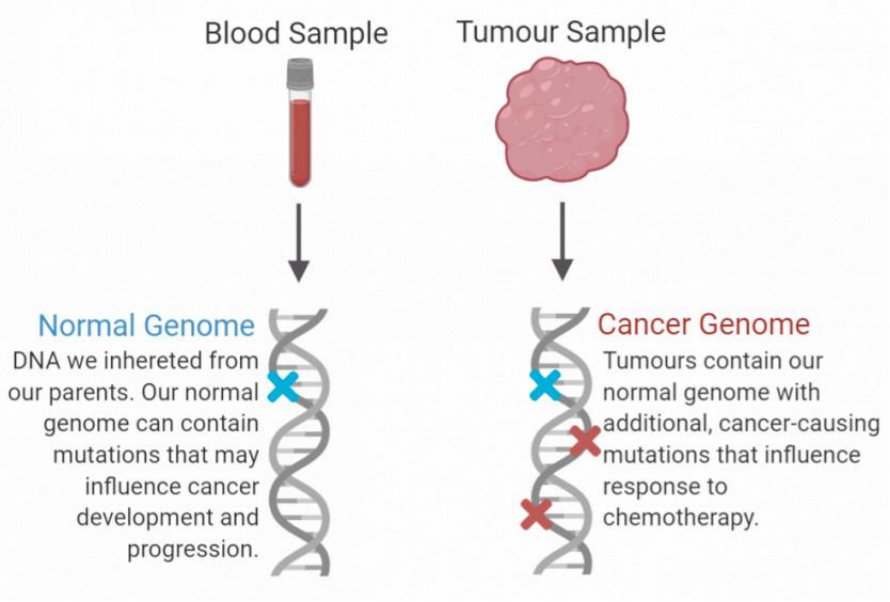
Cancer evolution
•Cancer cells “evolve” to out-compete other somatic cells.
•Cancer emphasizes the importance of the distinction between germline and soma.
•Cancer arises (typically) as the result of multiple mutations. Therefore, both contingency and chance are important.

Most cancers are caused by **multiple somatic mutations** that transform normal cells into tumors.

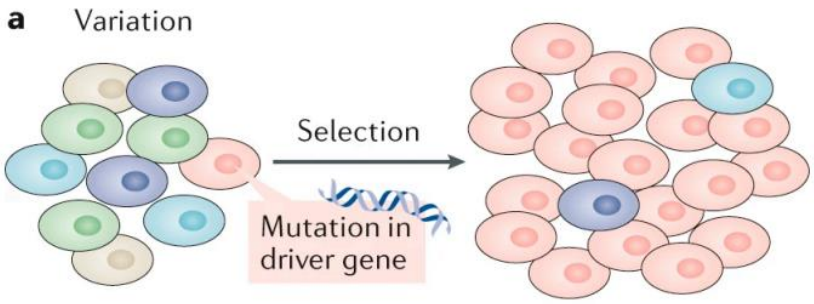
Cancer is not transmitted from one generation to the next, but **predisposition** to getting certain kinds of cancer often is.

Mutations arise during cell division — and reveal cell lineages
Cancer mortality rates differ between mammalian taxa.

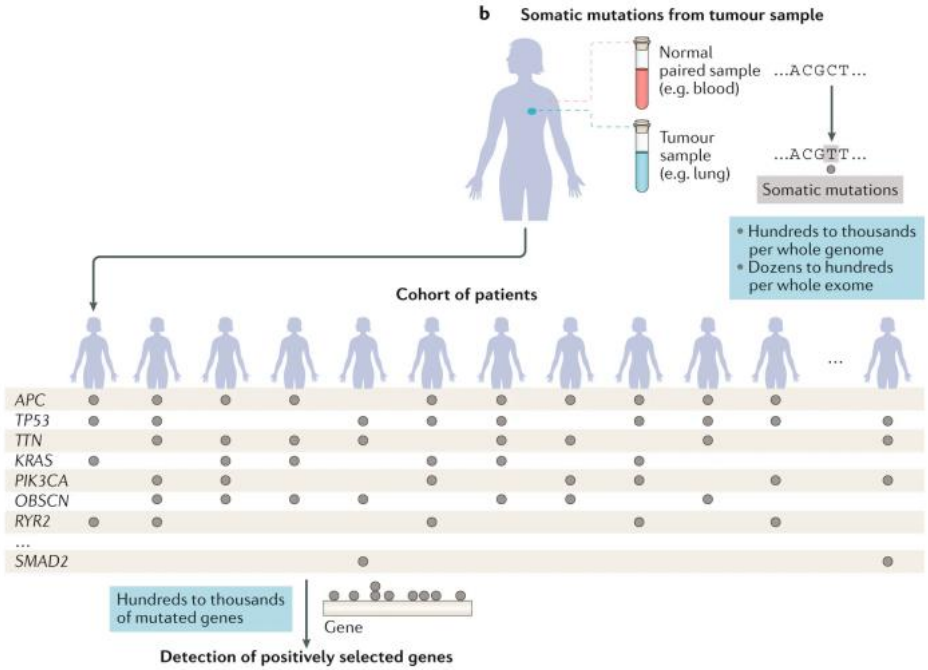
Given ongoing incidents of somatic mutation, how do we differentiate cancer-associated mutations from other mutations?



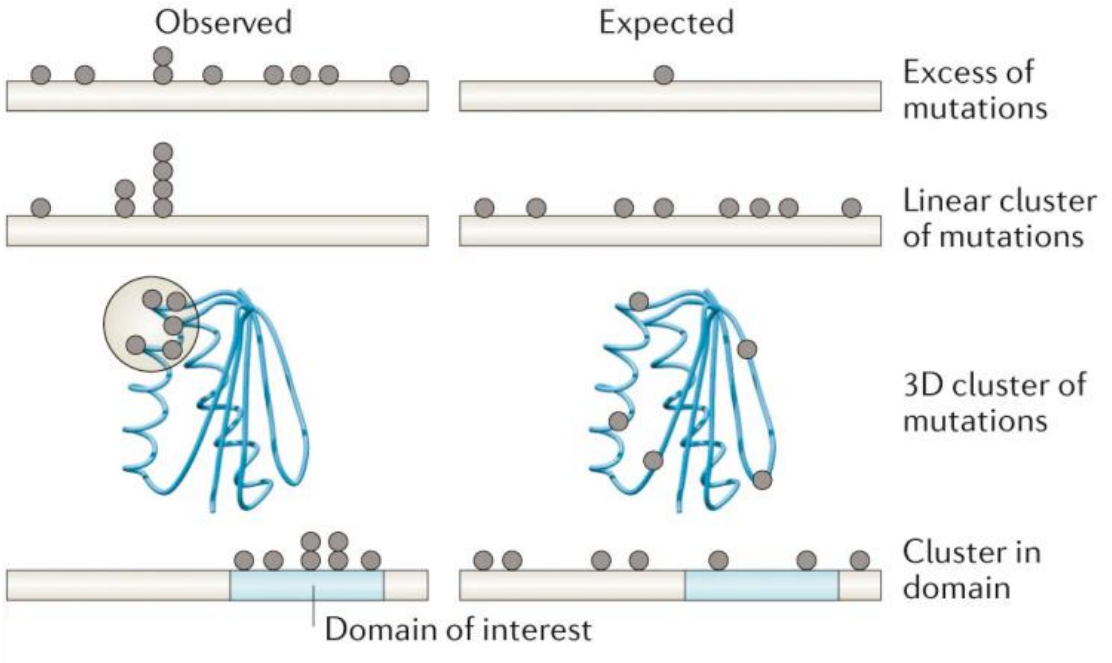
Driver mutations allow uncontrolled cell proliferation



How to detect driver mutations - look for **positive selection**



Signals of positive selection



How do we use cancer genomics to improve cancer diagnosis and treatment?

• Improved understanding of the molecular bases of cancer progression

• Improved understanding of the differences between cancers

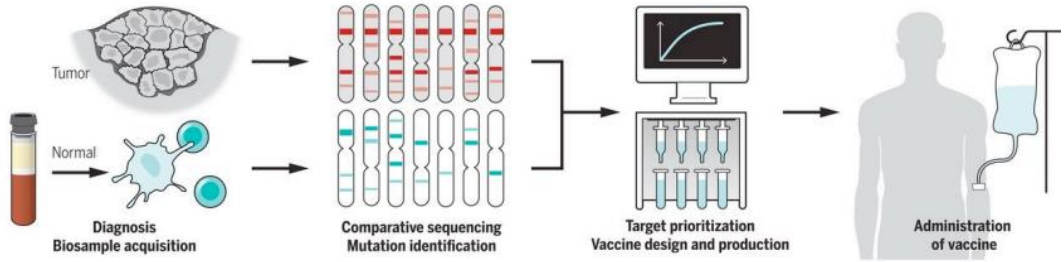
• Possibility of improving cancer treatment by tailoring drug application to the underlying specific mutant molecules involved

Chimeric Antigen Receptor (CAR) T-cell Therapy

1. Remove blood from patient to get T-cells
2. Make CAR-T cells in the lab (by inserting gene for CAR)
3. Grow millions of CART T cells
4. Infuse CAR T cells into patient
5. CAR T cells bind to cancer cells and kill them.

Limitations of CAR T-cell Therapy
1) Cancers are diverse and caused by mutations in many different driver genes
2) Vaccine development is slow and expensive

Development of mRNA vaccines for cancer immunotherapy



Personalized medicine gives more patients access to effective treatment.

Genetics of Infectious Disease

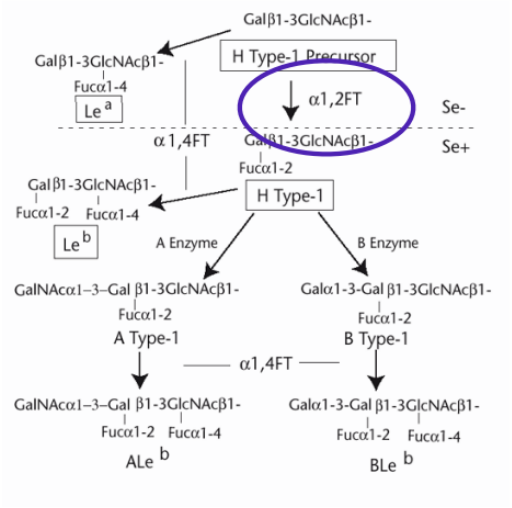
Best evidence for role of infection in human history comes from evolution.

Families of genes with strongest signals of historical positive selection along human lineage:
- Defense against infection and immunity
- Homeotic transcription factors (development)
- Immunoglobulin receptor family member

Susceptibility to infection is highly heritable with strong genetic component.

The genetics of infectious disease is hard to study. Don't know if people didn't get disease because they aren't susceptible or because they weren't exposed

Norovirus: positive sense, single segment RNA virus
Noroviruses are highly contagious. But researchers in Europe noticed that even in environments where there was universal exposure (cruise ships, daycare centers) 20% of people didn't get sick. 20% of native Europeans are "**H antigen non-secretors**": Are they the resistant population?



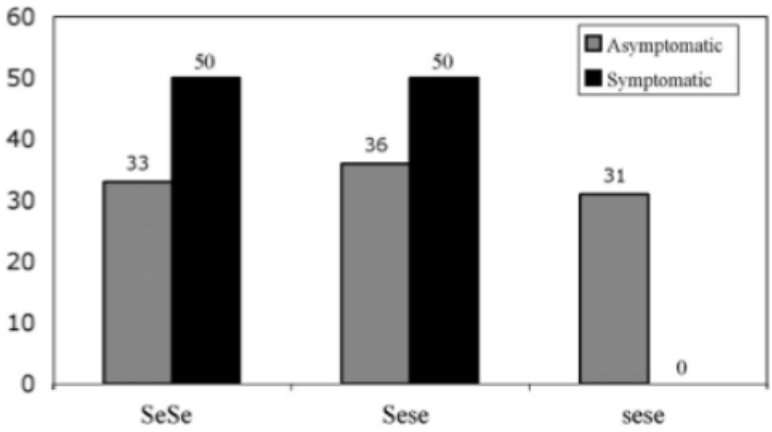
FUT1/FUT2

required for H antigen production

FUT1 in RBCs (and elsewhere)

FUT2 in epithelia

Genetics of Norovirus resistance



A Homozygous Nonsense Mutation (428G→A) in the Human Secretor (*FUT2*) Gene Provides Resistance to Symptomatic Norovirus (GGII) Infections

Resistance comes at a cost, non-secretor phenotype associated with altered risk of bacterial and influenza infection, as well as atypical gastric microbiota and increased risk of miscarriage

So study people who were almost certainly exposed...

Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposures

HIV resistance allele is 32bp deletion in CCR5 gene.

Long-Term Control of HIV by *CCR5* Delta32/Delta32 Stem-Cell Transplantation

Two general types of polymorphisms can affect susceptibility to infectious disease
1. Polymorphisms that affect cellular processes used by pathogen to enter cells and replicate
2. Polymorphisms in genes involved in immunity

Lots of immune genes linked to pathogens
MHC are antigen presenting proteins. MHC proteins present antigens to other cells, who decide what to do. MHC Complex has many **HLA** genes, HLA loci are massively polymorphic, almost certainly the result of selection to diversify in face of complex pathogen threat.
HLA allele frequencies vary among populations

Polymorphisms affect antigen presentation and receptor binding. T-cell recognition is **specific to HLA protein** and antigen. HLA diversity increases repertoire of pathogens recognized by population.
Mutations creating new HLA alleles can reduce response to some pathogens, but also provide defense against pathogens that have evolved to evade common HLA alleles.

Both alleles at each HLA locus are expressed in each cell.

Specific HLA alleles associated with resistance/ susceptibility to specific diseases, which is the major source of polymorphic response to pathogens.

Infectious agents and vectors have genetics too, generating rapid accumulation of mutations to avoid immune surveillance.

Cell biology of Infectious Disease

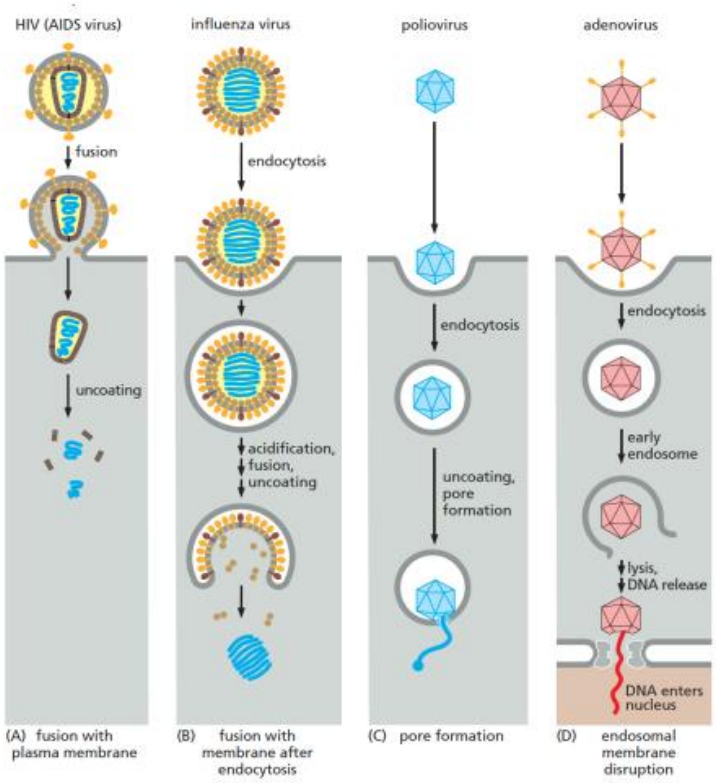
1.The global and historical burden of infectious disease

Infectious disease remains a huge contributor to human mortality.
Many infectious diseases cause extensive morbidity (illness) in addition to mortality.

2.A quick look at viral diversity

Viruses come in different shapes and sizes.

Viruses commandeer host cell machinery
1. Viruses enter cells by membrane fusion or endocytosis



2. The infectious particle is disassembled

3. Viral genomes (RNA or DNA) are replicated

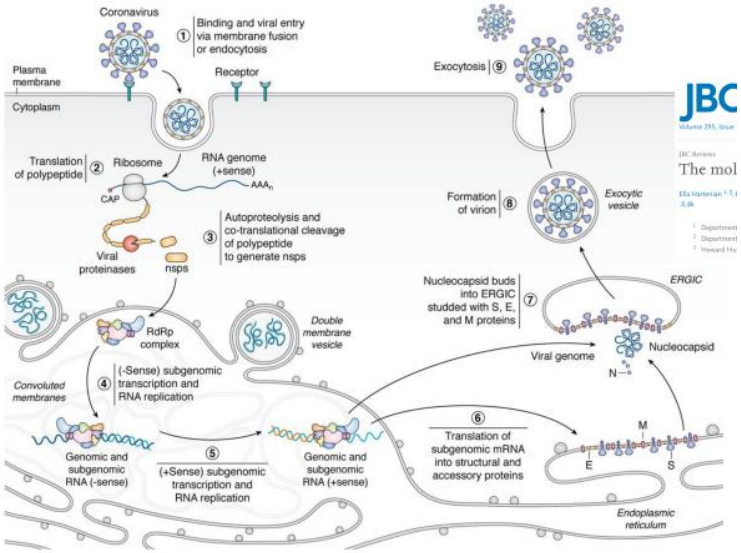
4. Viral genes are transcribed and proteins translated (note: RNA (+) strand viruses can start this immediately

5. New virus particles (virions) assemble

6. Virions are released from cells

3.SARS-CoV2 entry and exit mechanisms

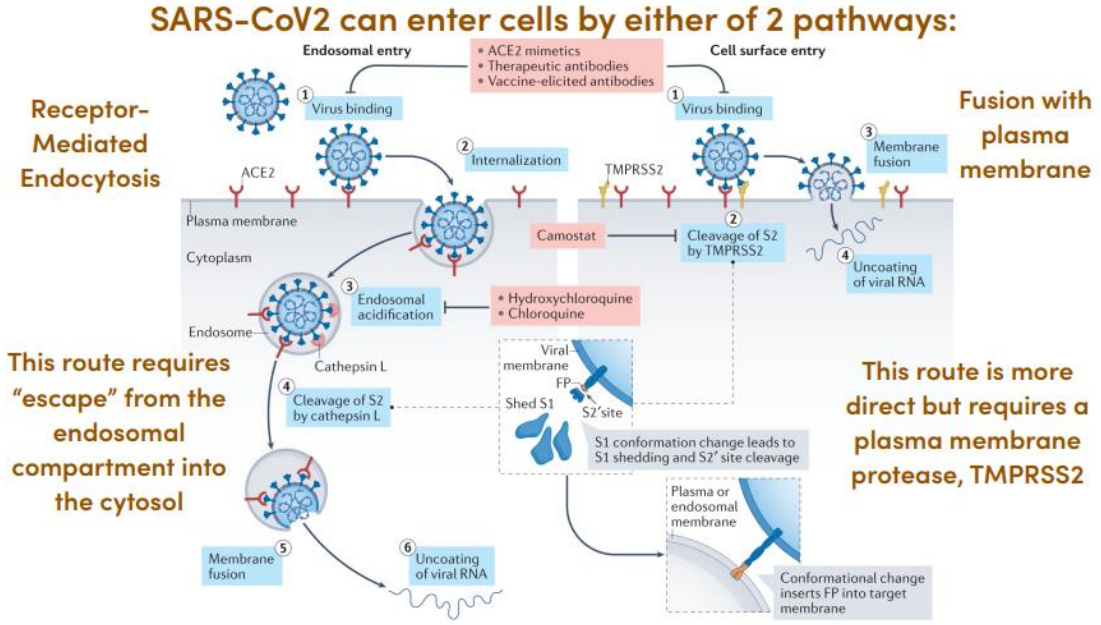
Intracellular life cycle of coronaviruses



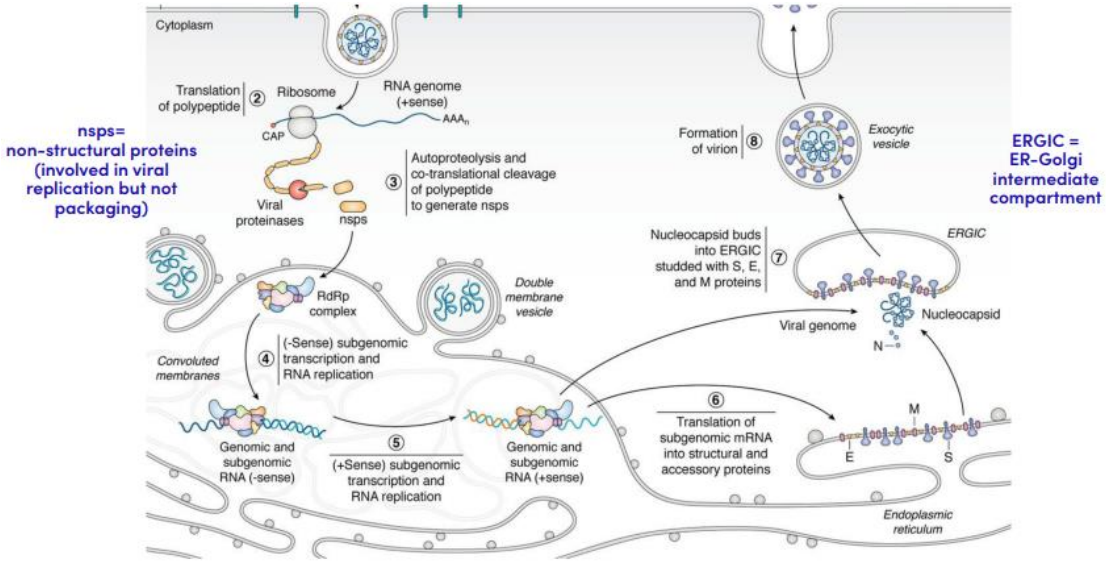
Some viruses enter cells via receptor-mediated endocytosis.

How do we know if a virus enters cells via receptor-mediated endocytosis? Receptors on the host cell for the Virus!

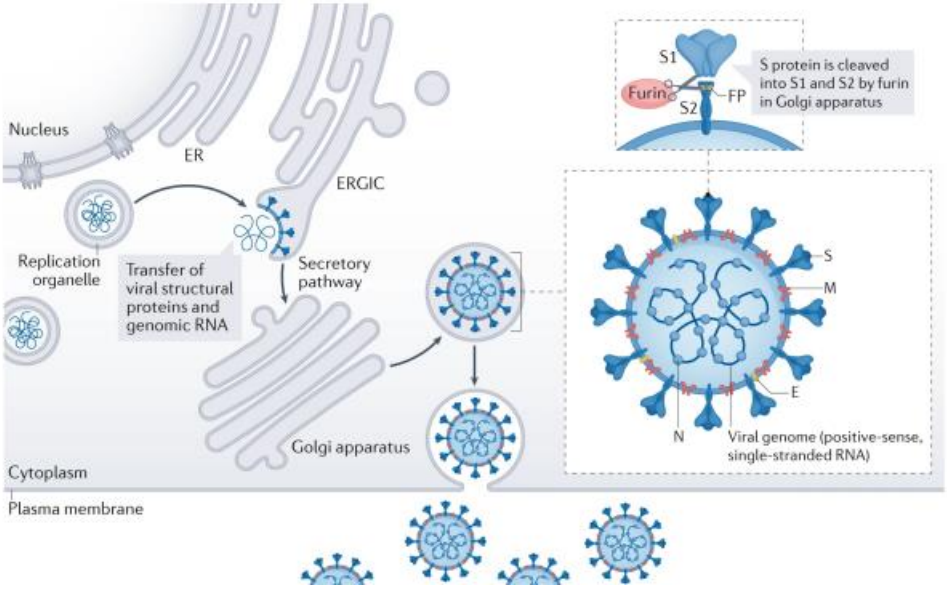
The receptor for SARS-CoV2 is a protein called angiotensin-converting enzyme 2 **(ACE2)**.

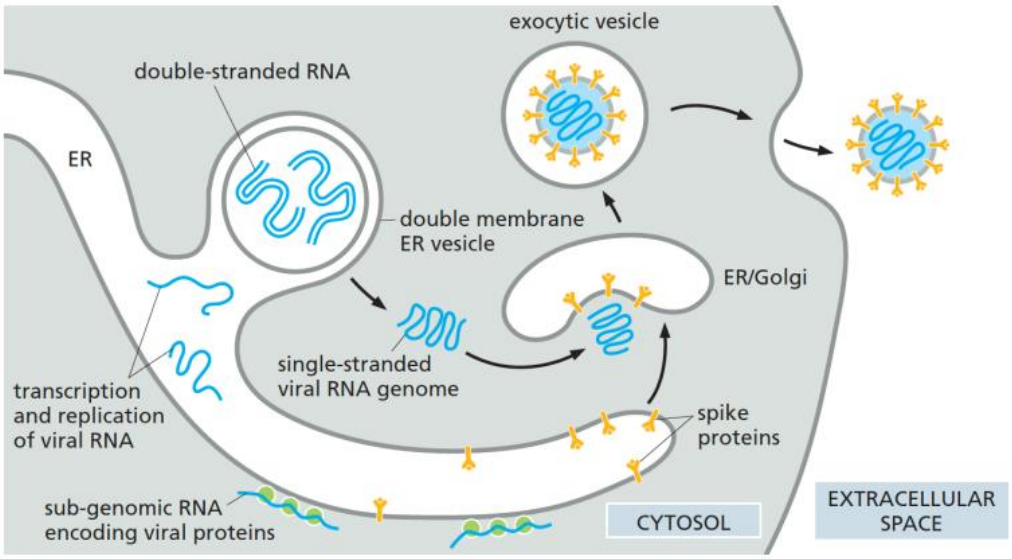


Once inside the cell, viral proteins are translated, and the replication process begins



Virus assembly and budding occur in the ERGIC (ER-Golgi Intermediate Compartment)

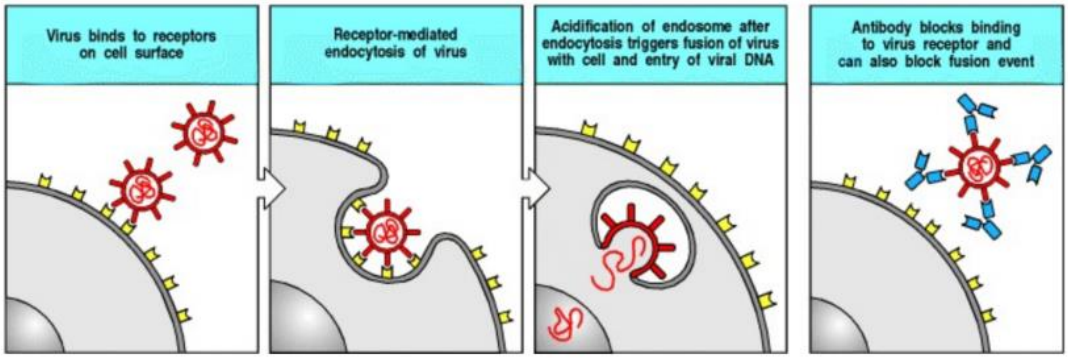




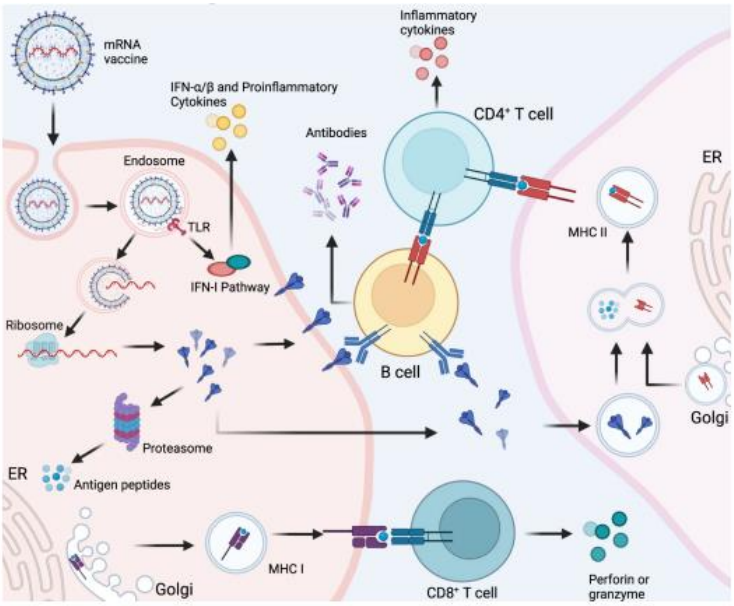
4.mRNA vaccines

Antibodies can recognize - and block - a variety of epitopes[表位]

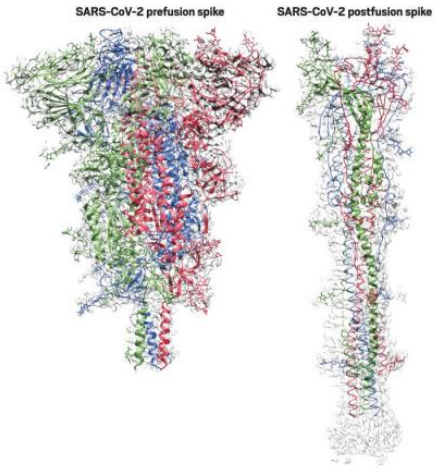
Injection of an “antigen” (usually a purified protein or protein fragment) into a mouse, guinea pig, rabbit, or human leads to production of antibodies that bind tightly to that molecule. Antibodies can “neutralize” viruses by blocking binding to cell surface receptors.



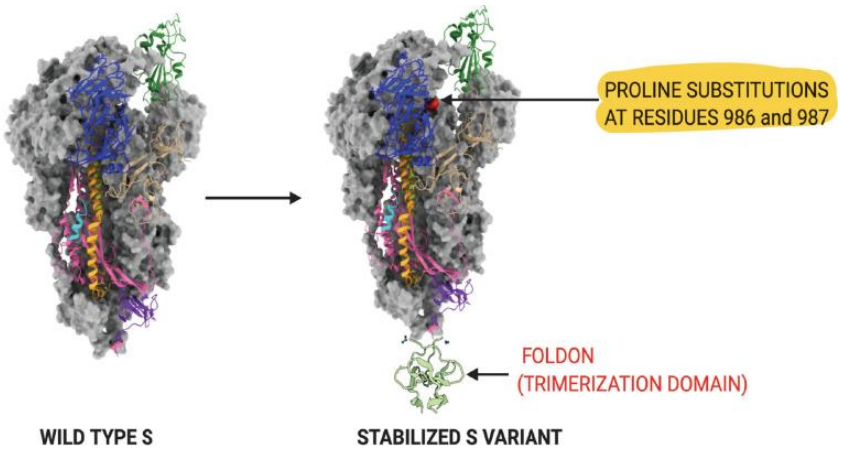
mRNA vaccines lead to **antigen production** by our own cells.
Delivery of mRNA molecules into cells is a crucial aspect of the vaccine.



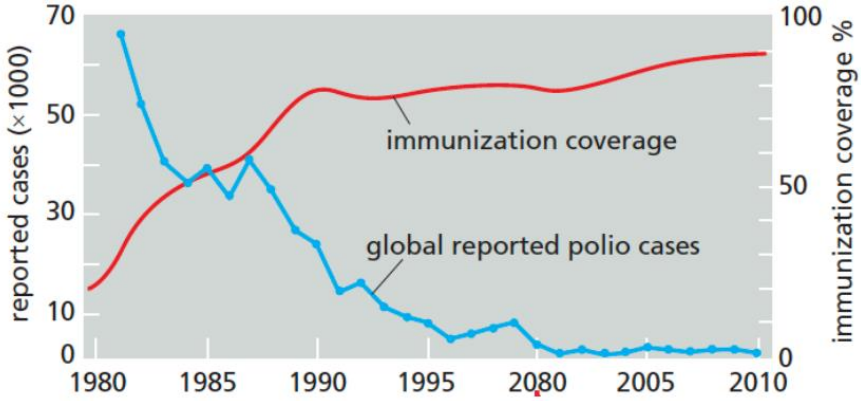
The spike protein is flexible - that's part of it's job



Stabilization of the "pre-fusion" conformation by amino acid substitutions helps to elicit neutralizing antibodies

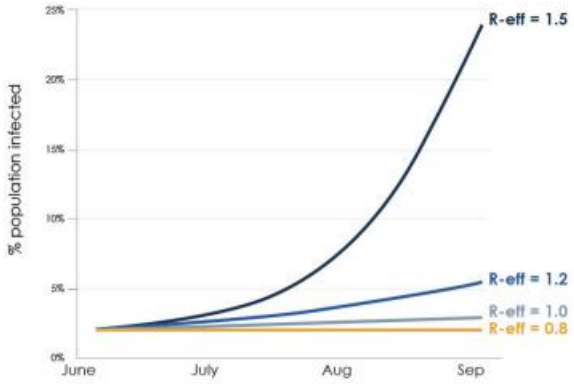


Vaccines can eradicate infectious diseases



Epidemiology and Molecular Virology of SARS CoV-2

R effective: If the R effective is greater than 1, COVID-19 will spread exponentially. If R effective is less than 1, COVID-19 will spread more slowly and cases will decline. The higher the value of R effective, the faster an epidemic will progress. The following graph illustrates the change in growth as R effective increases.



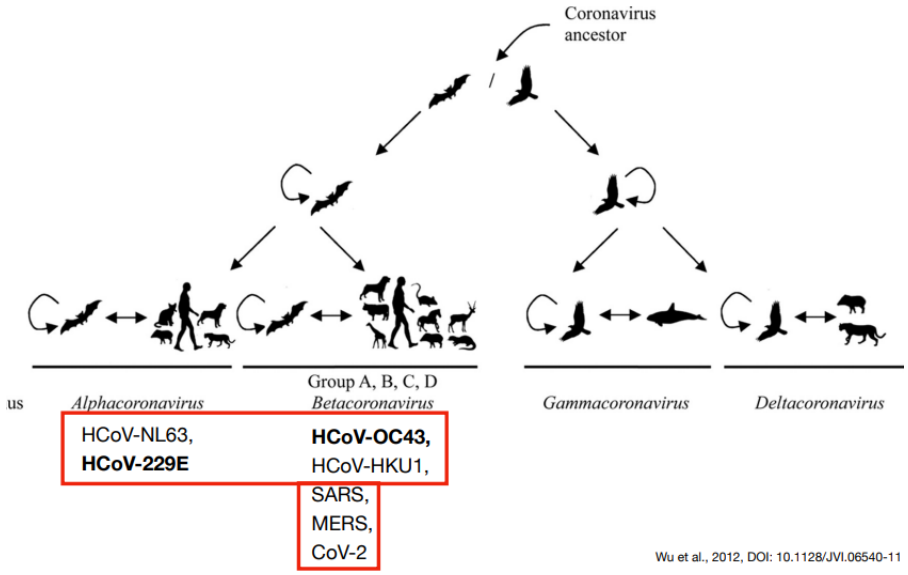
Adapted from CEBM

use what you've learned in this class to be a critical thinker/ consumer of science information

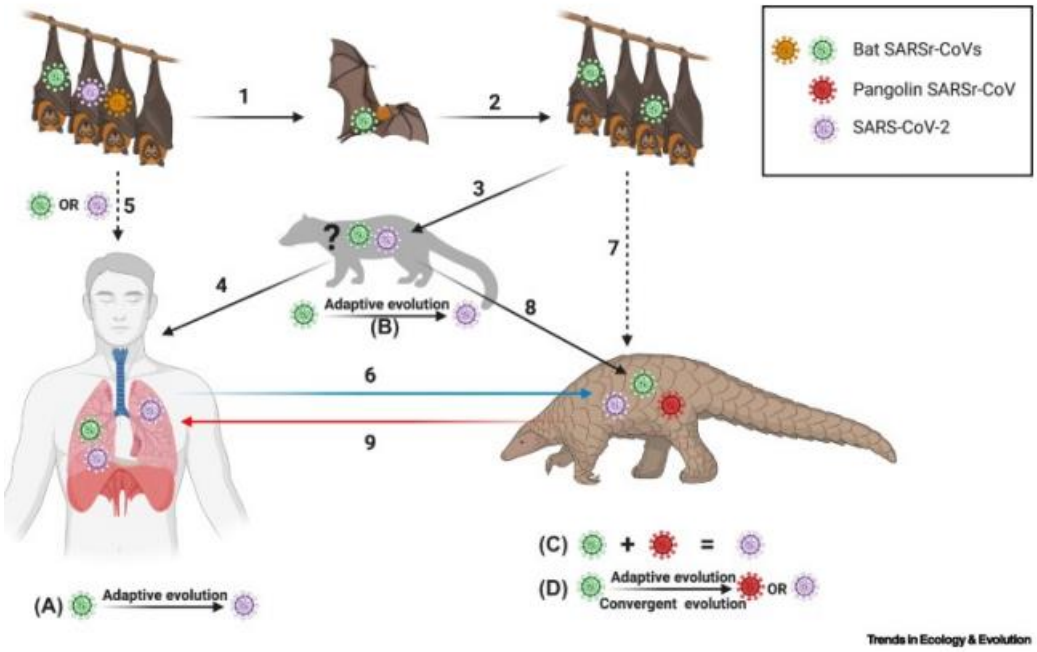
Coronavirus facts:
ssRNA (+sense), 32kb, largest known RNA virus genome, mutate and change at a high rate.

Terminology
- SARS: Severe Acute Respiratory Syndrome
  - The current pandemic is caused by the **SARS CoV-2** virus
  - The diseased caused by SARS CoV-2 is called **COVID-19**
- RdRP: RNA-dependent RNA polymerase

There are 7 human CoVs, present in the alpha-and betacoronavirus genera



Possible origins of SARS CoV-2

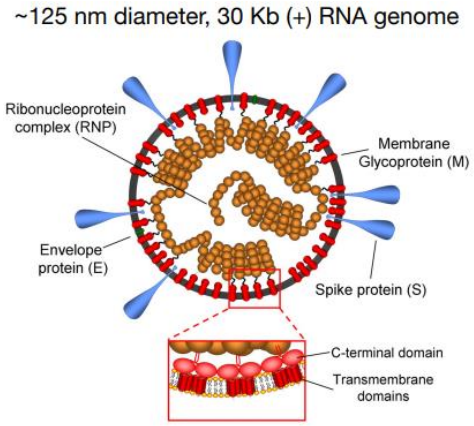


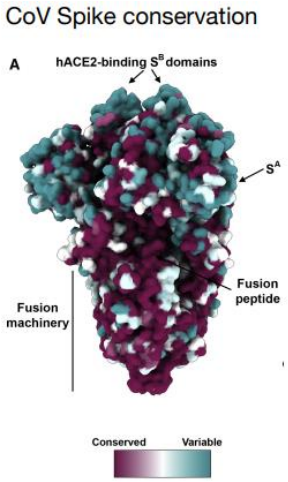
How would you investigate the origin of SARS CoV2?
Using genomics for epidemiology!
Nextstrain / ncov / gisaid / global / 6m

Why has CoV-2 been so much harder to control than SARS?

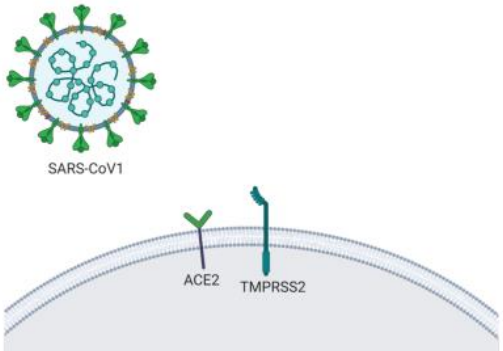
CoV-1/SARS	CoV-2/COVID19
'spillover' reservoir known: civet cats *-cull to break the chain*	'spillover' reservoir unknown
Most transmission occurred in hospital setting (hubs) *-implemented barrier nursing*	Widespread community transmission
No transmission until 24-36h after symptoms, lack of asymptomatic cases *-contact tracing effective*	Abundant asymptomatic/mild cases

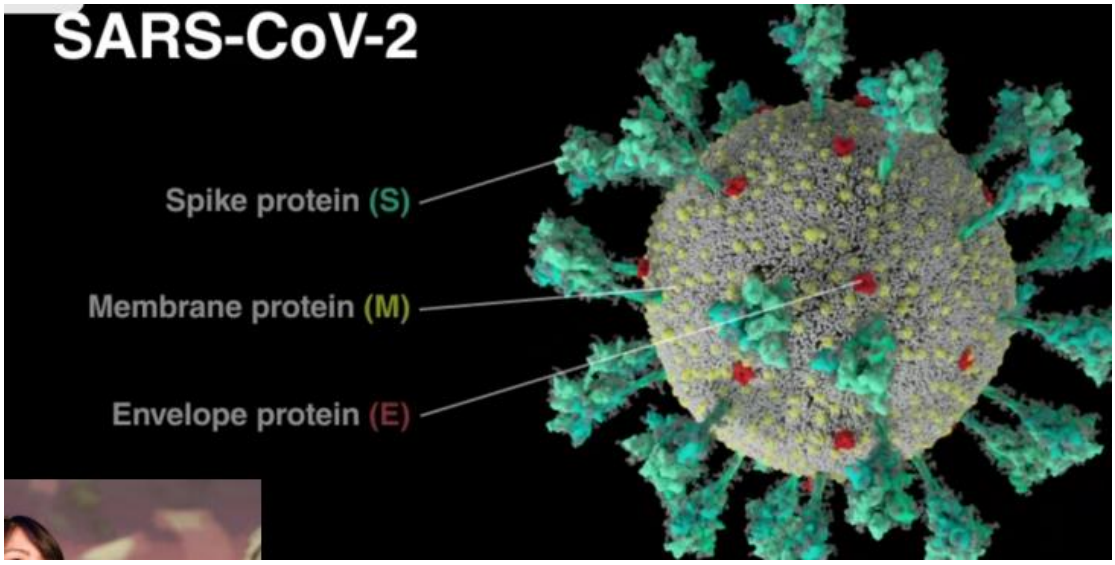
Zooming in on the CoV2 viral particle



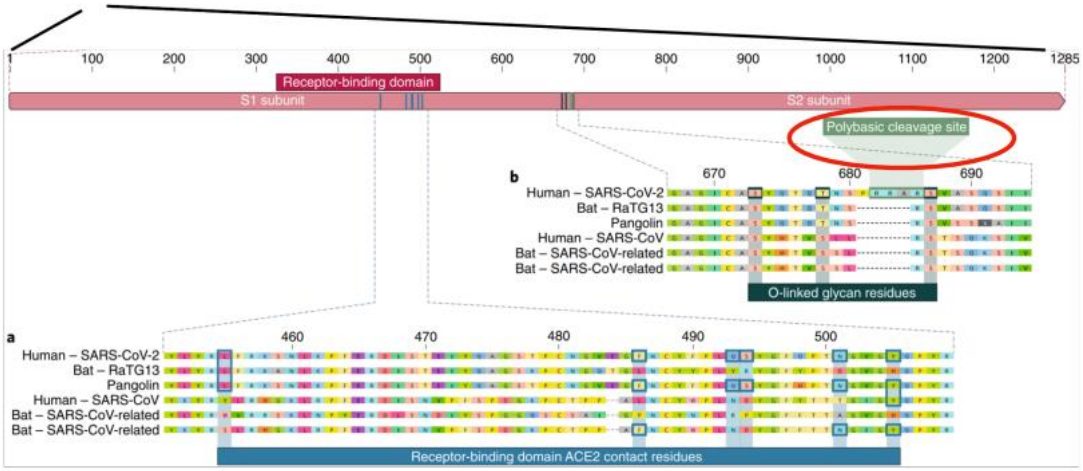


CoV-2 entry is driven by interactions between Spike and angiotensin-converting enzyme 2 (ACE2); subsequent protease cleavage drives fusion



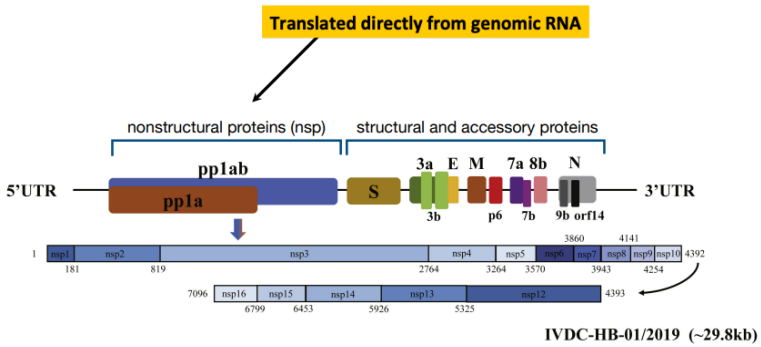


Acquisition of polybasic cleavage site in CoV-2 spike may increase viral transmissibility

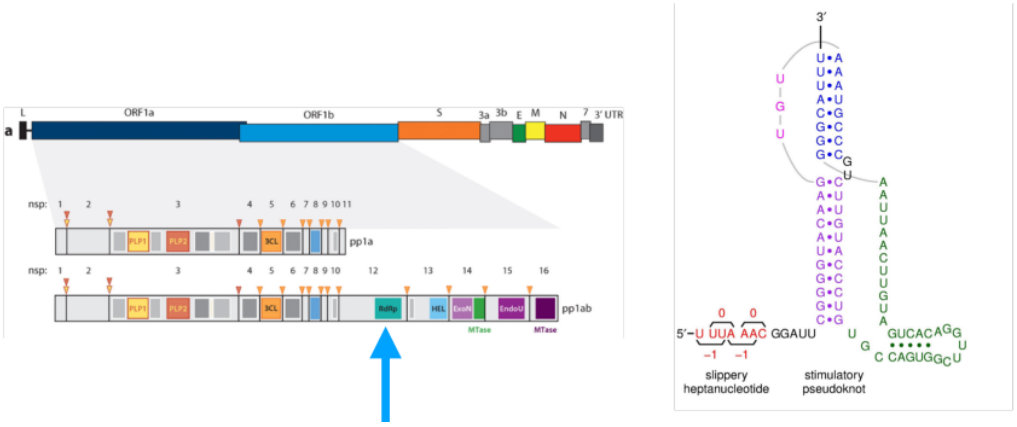


The RNA genome has to abide by eukaryotic "rules" of gene expression and translation

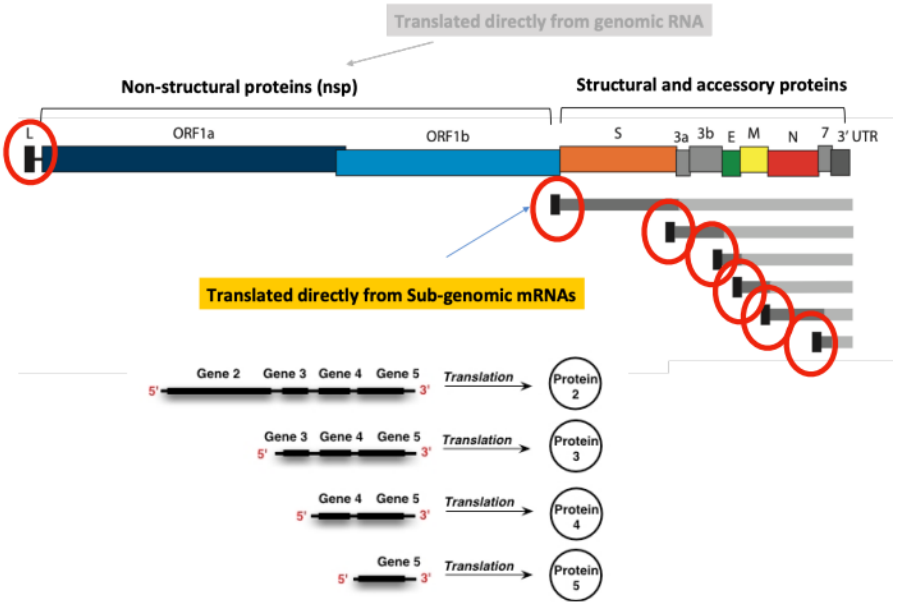
The 2019-nCoV genome was annotated to possess ~14 ORFs encoding 27 proteins



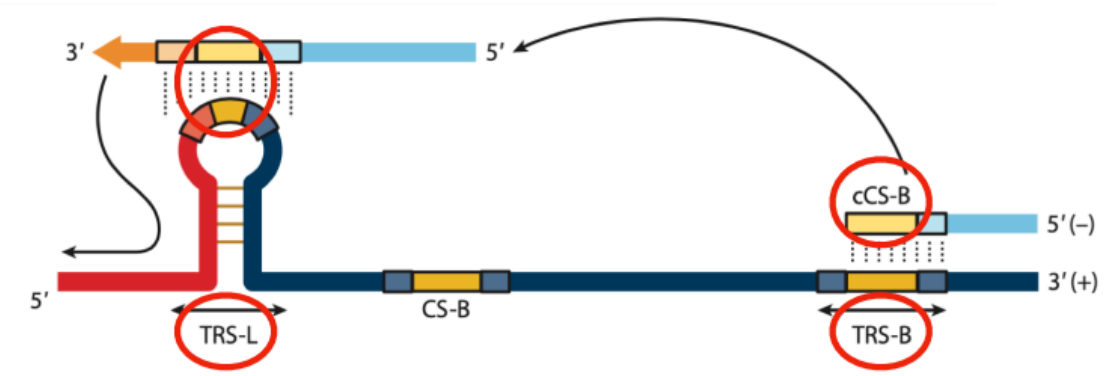
Programmed ribosomal frameshifting generates two polyproteins encoding the replicase proteins.



Structural proteins are made from a nested set of sub-genomic mRNAs with shared 5’ and 3’ sequences.



Sub-genomic RNA transcription is discontinuous and is facilitated by shared transcription regulatory sequences



The CoV replicase requires functional integration of RNA polymerase, capping, and proofreading activities

Genomics is accelerating the end of the pandemic

•Genome sequencing allowed rapid identification of the causative virus

•The genome sequence revealed that CoV2 was related to known and well-characterized viruses

•…this allowed scientists to quickly infer the mechanism of viral entry into host cells

•…and this allowed pharma to quickly develop effective vaccines

•Genomics also allows us to detect new variants and observe how they move through populations