

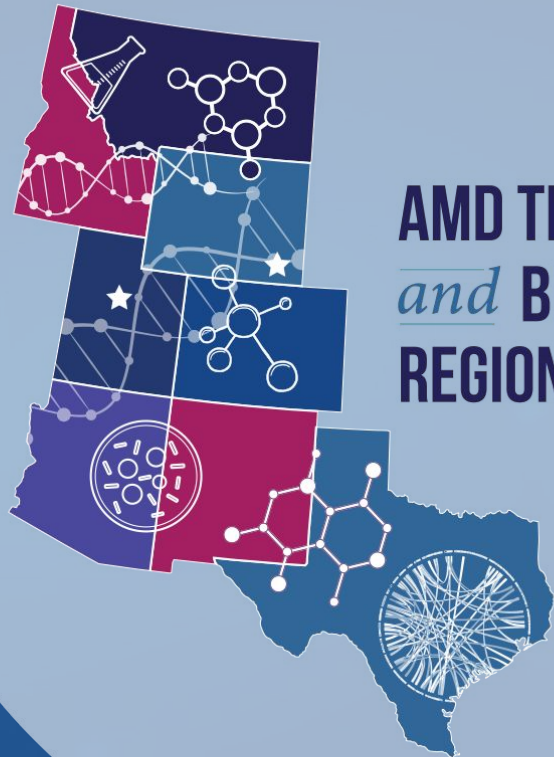
MTN Region WFD 2020 Virtual Training: Session 2 - Background on Viral Genomics and Coronaviruses

Joel R. Sevinsky, Ph.D.

WY PHL AMD Training Lead

6 July 2020 - 24 July 2020

joel@sevinsky.com



**AMD TRAINING LEAD
and BIOINFORMATICS
REGIONAL RESOURCE**

Housekeeping

Register for a GISAID account:

<https://www.gisaid.org/registration/register/>



What Most of Us Know

PulseNet

- Isolate a bacterial colony
- Grow it up
- Extract genomic DNA
- Prepare a library with DNA Flex
- Run library on MiSeq
- Process data in Bionumerics



Bacterial Pathogen Bioinformatics

Predominantly gene/genome based characterization

- Species
- Serotype
- AMR genes
- Virulence genes

We use SNPs/alleles for phylogenetics and epidemiology.

Do we ever talk about where the SNPs are?

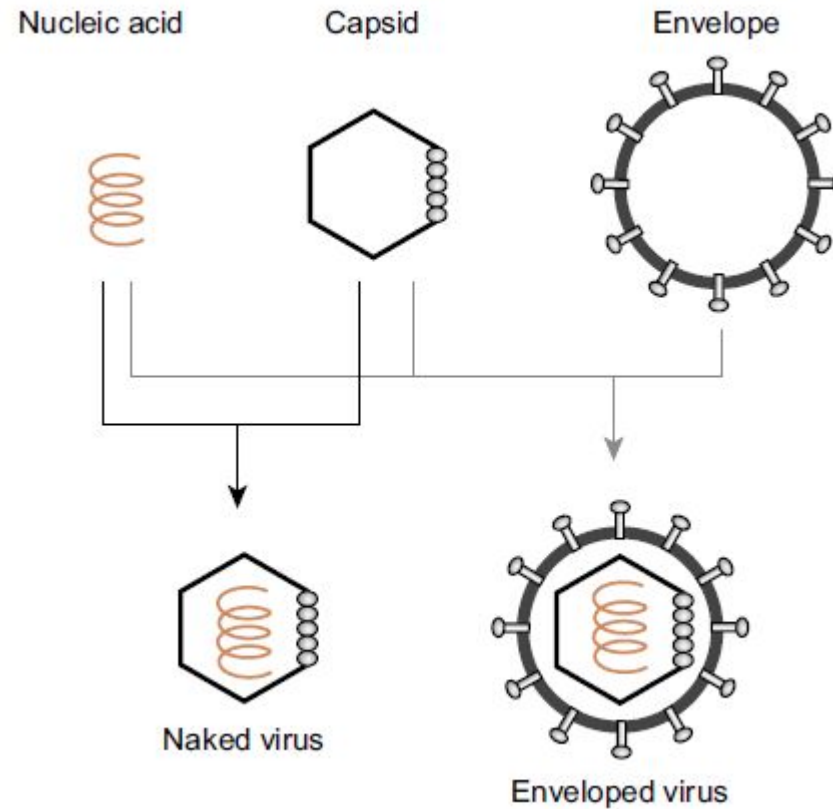


NGS: Bacteria vs Viruses

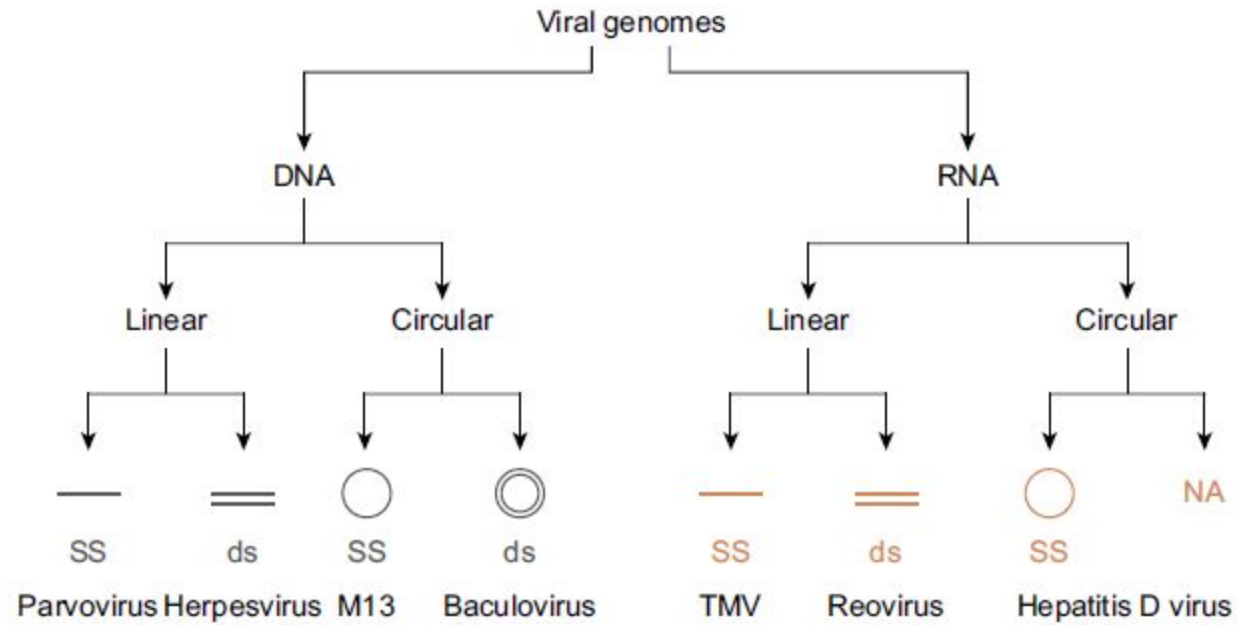
	Bacteria (enterics)	Viruses (coronaviruses)
Genome size	0.5 - 12 Mb (~1.6 - 5.4 Mb)	2 - 2,000 kB (~30 kb)
# of genes	High (~1,600-5,000)	Low (dozens)
Rate of mutation	Low	High
Nucleic Acid	dsDNA	dsDNA, dsRNA, ssDNA, ssRNA
Sequencing method	WGS mostly	Metagenomics, amplicon



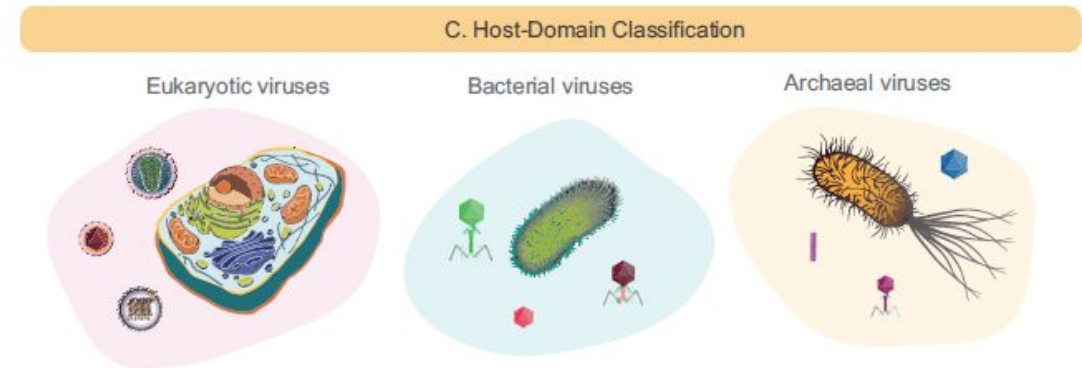
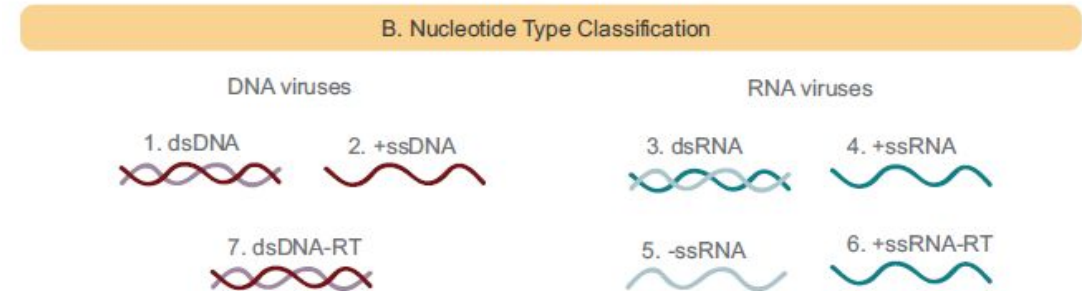
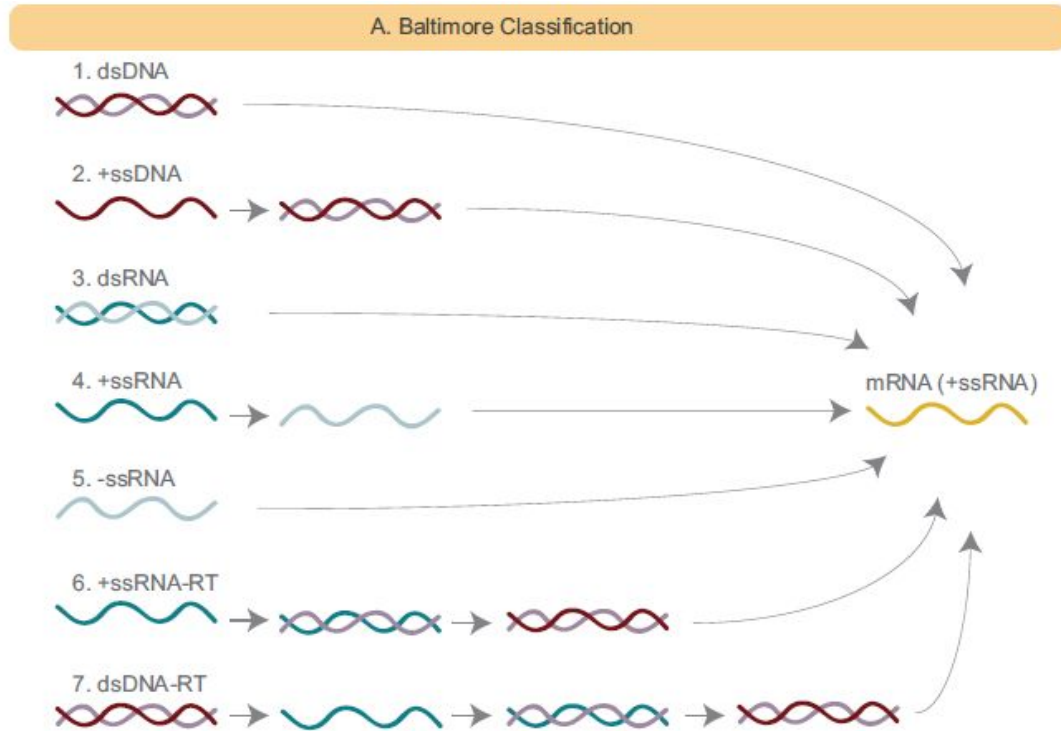
Overly Brief Intro to Viral Pathogens - Structure



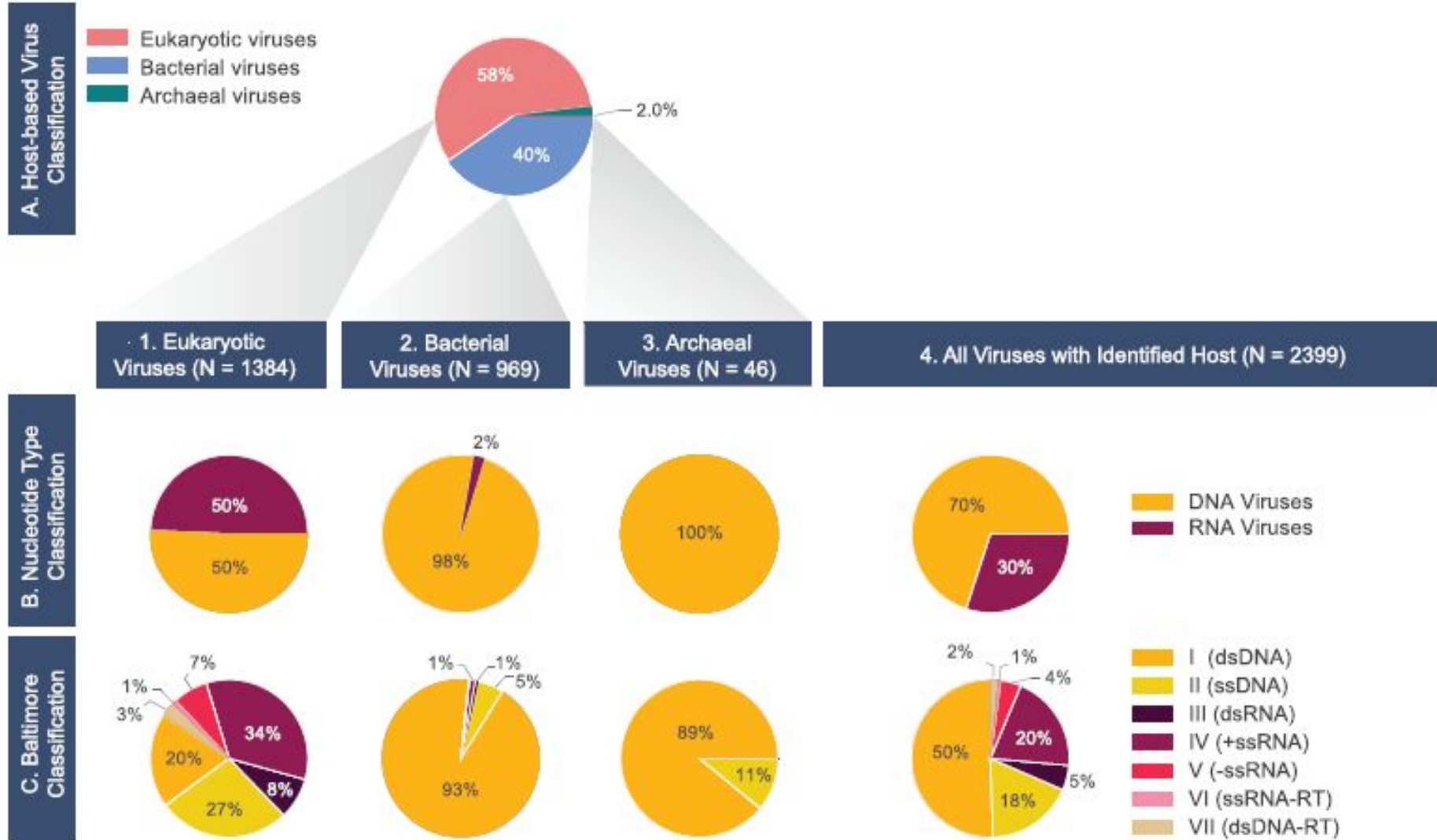
Viral Genomes



Classification



Diversity



Replication

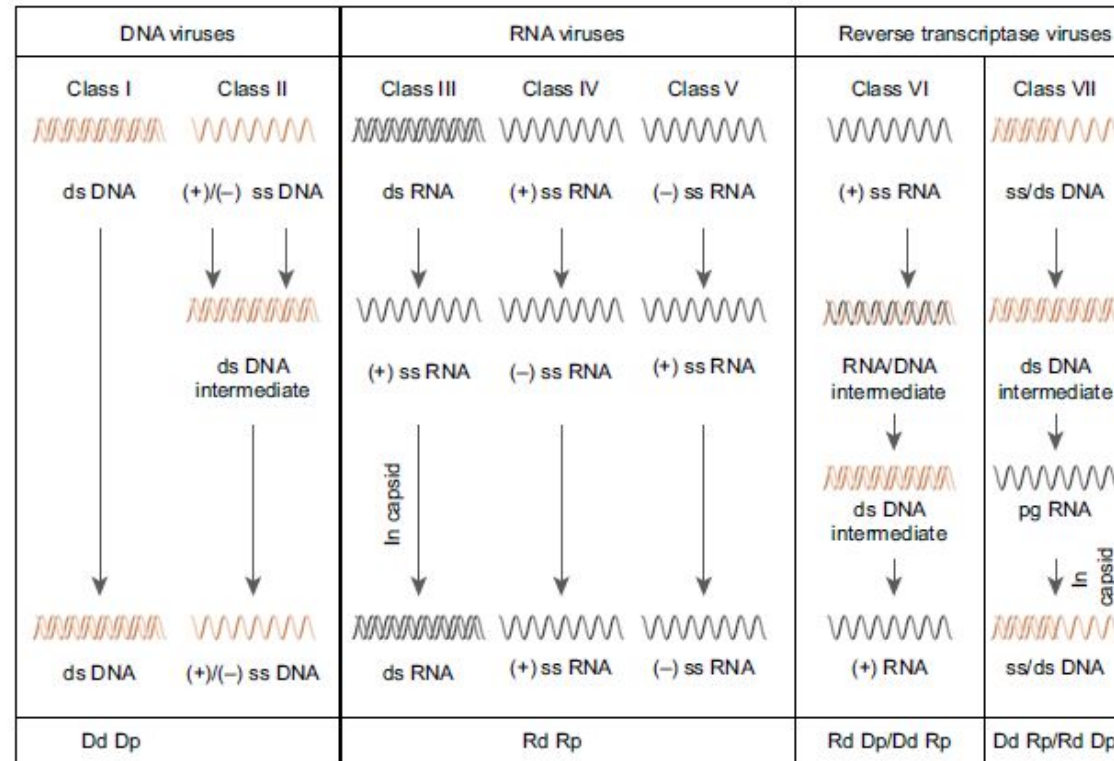
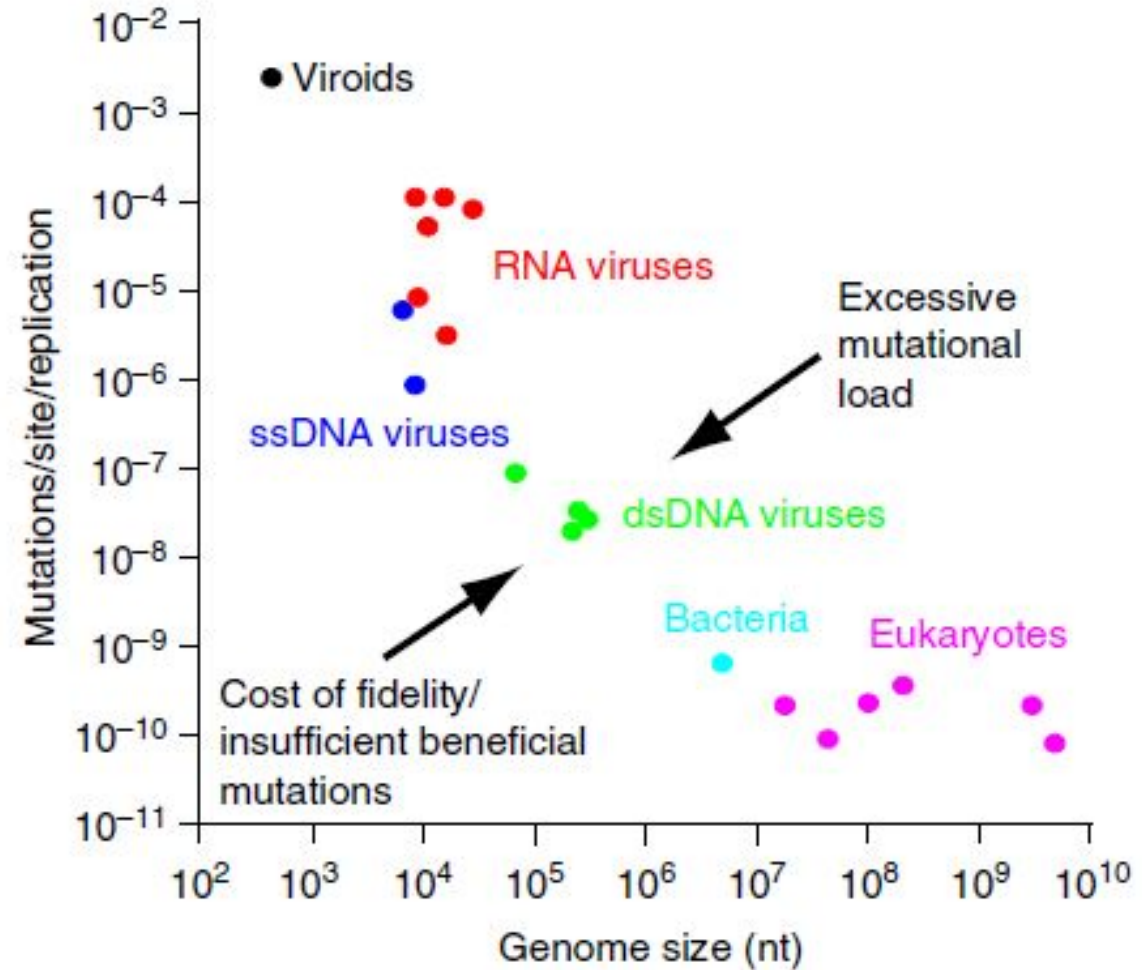


FIGURE 1.12 Mechanisms of viral genome replication. Baltimore classification of viruses and their modes of replication and generation of viral mRNA. DNA genomes are depicted with *black helices*, whereas RNA genomes are depicted with *orange helices*. The enzymes replicating each class genome are abbreviated as follows: *DdDp*, DNA-dependent DNA polymerase; *RdDp*, RNA-dependent DNA polymerase or reverse transcriptase; *RdRp*, RNA-dependent RNA polymerase.



Mutation

Figure 3 The fundamental relationship between log mutation rate (mutations/site/replication) and log genome size in diverse organisms, including RNA viruses and viroids. The evolutionary forces that might be responsible for the observed range of genome sizes and mutation rates are shown. Modified from Holmes, E.C., 2011. What does virus evolution tell us about virus origins? *Journal of Virology* 85, 5247–5251 and Gago, S., Elena, S.F., Flores, R., *et al.*, 2009. Extremely high mutation rate of a hammerhead viroid. *Science* 323, 1308.



mRNA Synthesis

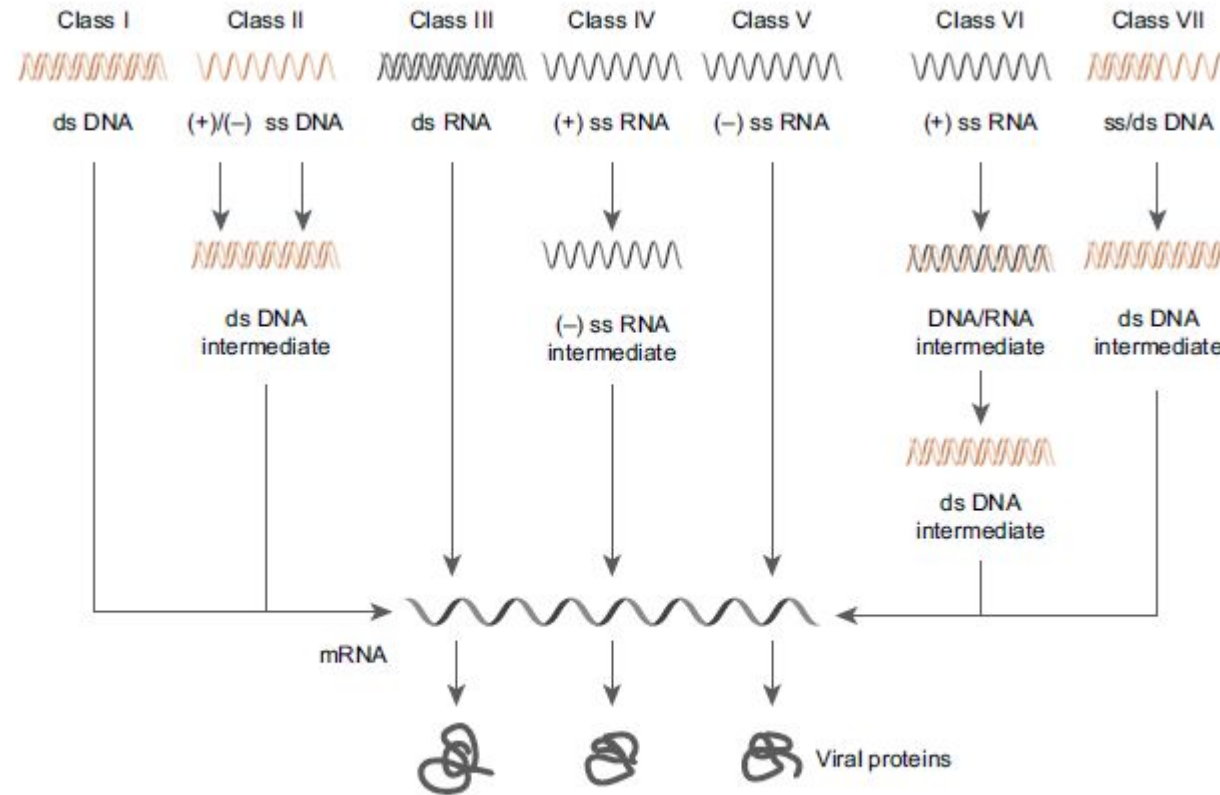
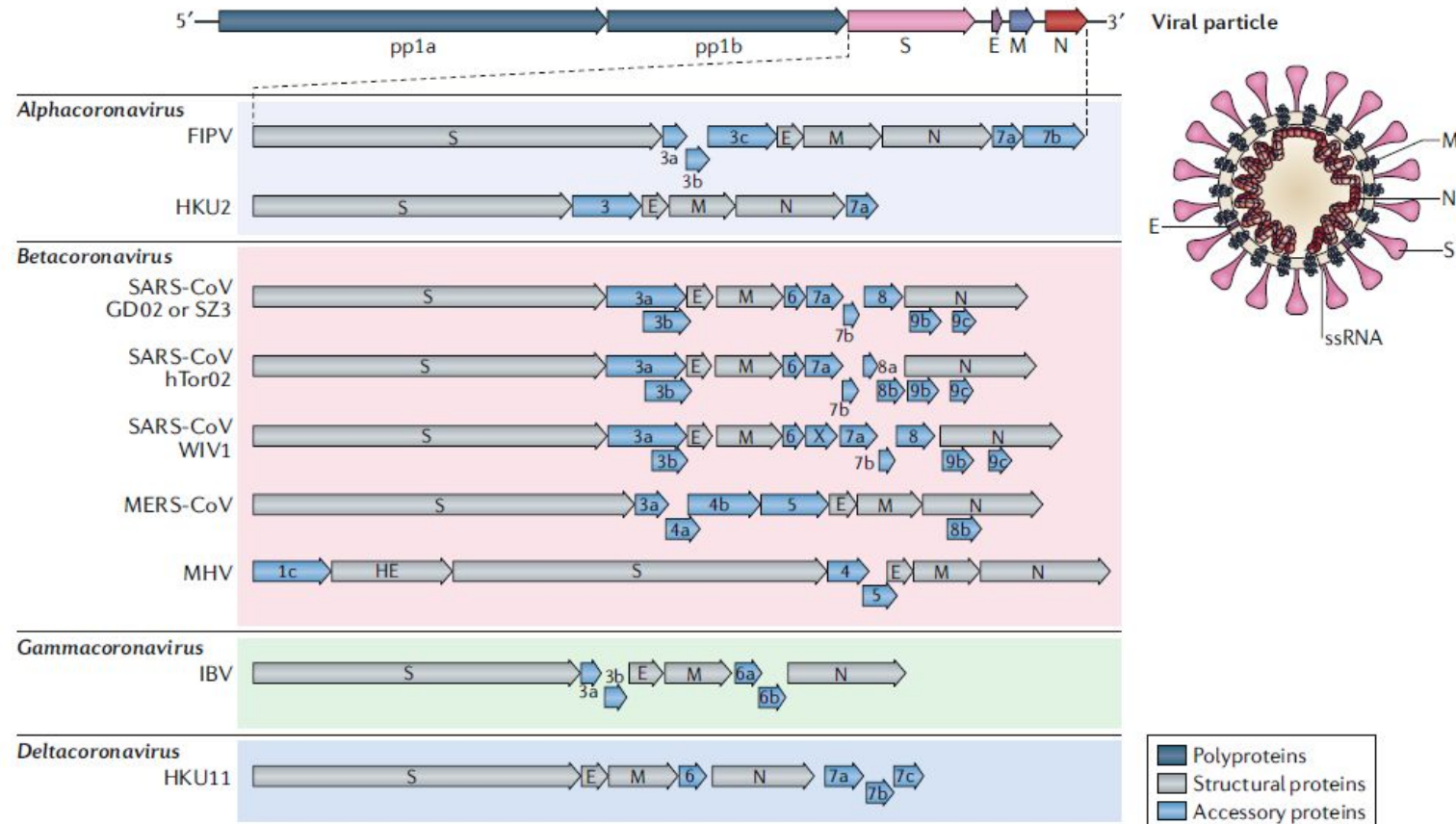


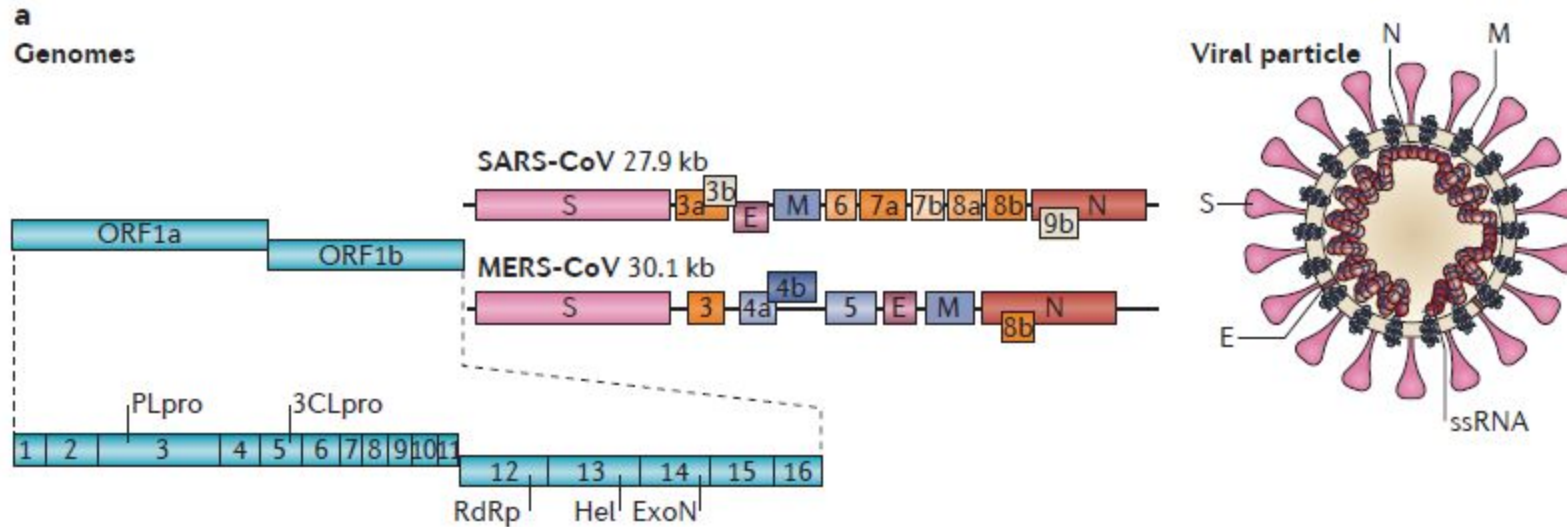
FIGURE 1.13 Mechanisms of viral mRNA synthesis. Baltimore classification of viruses and their mechanisms of mRNA generation. DNA is depicted with *black helices*, whereas RNA is depicted with *orange helices*.



Coronavirus Genome Structure



SARS and MERS

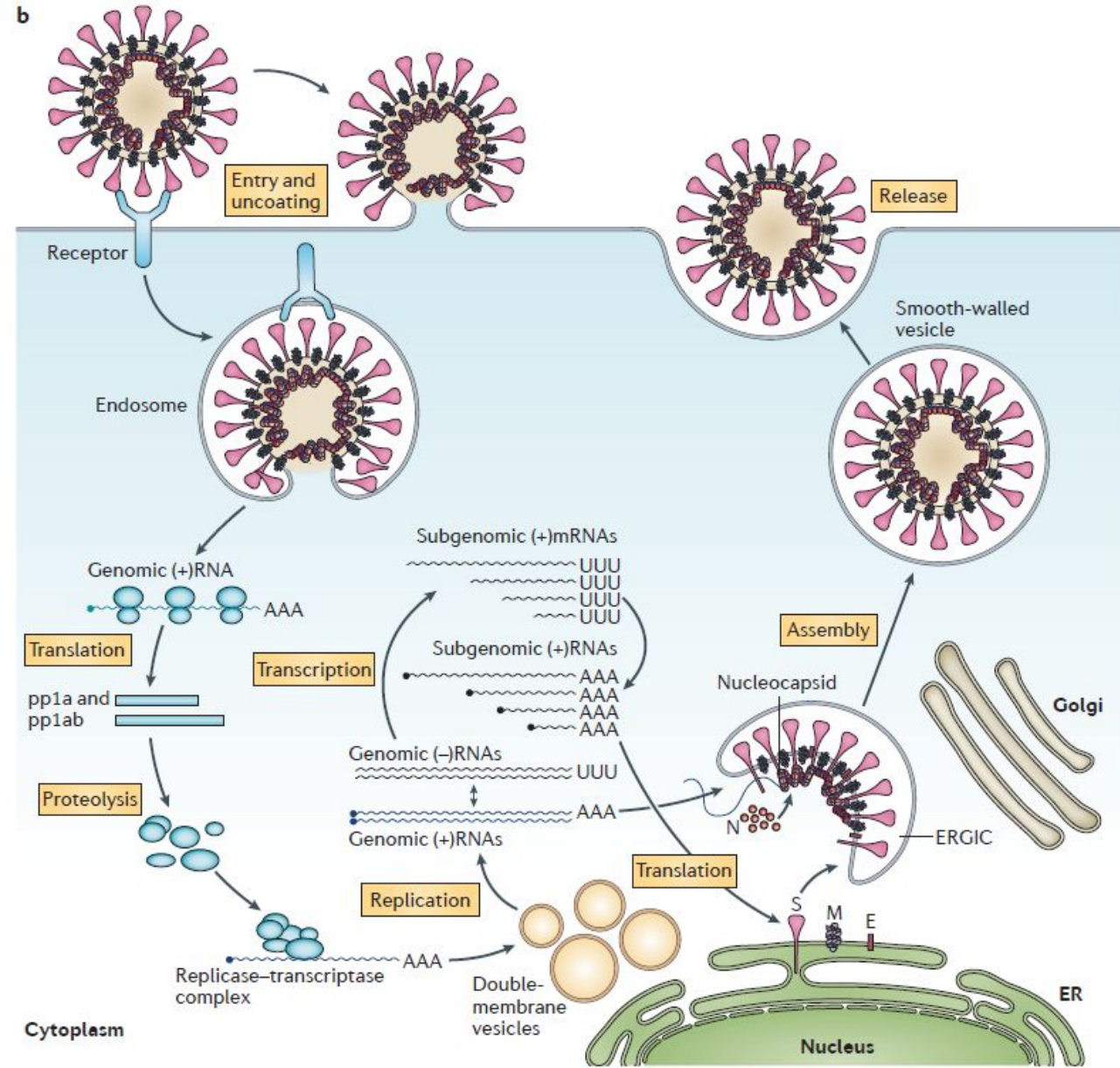


Coronavirus Genome Structure

Protein	Function	References
nsp1	Promotes cellular mRNA degradation and blocks host cell translation, results in blocking innate immune response	[125–128]
nsp2	No known function, binds to prohibitin proteins	[129, 130]
nsp3	Large, multi-domain transmembrane protein, activities include: <ul style="list-style-type: none"> • Ubl1 and Ac domains, interact with N protein • ADRP activity, promotes cytokine expression • PLPro/Deubiquitinase domain, cleaves viral polyprotein and blocks host innate immune response • Ubl2, NAB, G2M, SUD, Y domains, unknown functions 	[131–138]
nsp4	Potential transmembrane scaffold protein, important for proper structure of DMVs	[139, 140]
nsp5	Mpro, cleaves viral polyprotein	[141]
nsp6	Potential transmembrane scaffold protein	[142]
nsp7	Forms hexadecameric complex with nsp8, may act as processivity clamp for RNA polymerase	[143]
nsp8	Forms hexadecameric complex with nsp7, may act as processivity clamp for RNA polymerase; may act as primase	[143, 144]
nsp9	RNA binding protein	[145]
nsp10	Cofactor for nsp16 and nsp14, forms heterodimer with both and stimulates ExoN and 2-O-MT activity	[146, 147]
nsp12	RdRp	[148]
nsp13	RNA helicase, 5' triphosphatase	[149, 150]
nsp14	N7 MTase and 3'-5' exoribonuclease, ExoN; N7 MTase adds 5' cap to viral RNAs, ExoN activity is important for proofreading of viral genome	[151–154]
nsp15	Viral endoribonuclease, NendoU	[155, 156]
nsp16	2'-O-MT; shields viral RNA from MDA5 recognition	[157, 158]



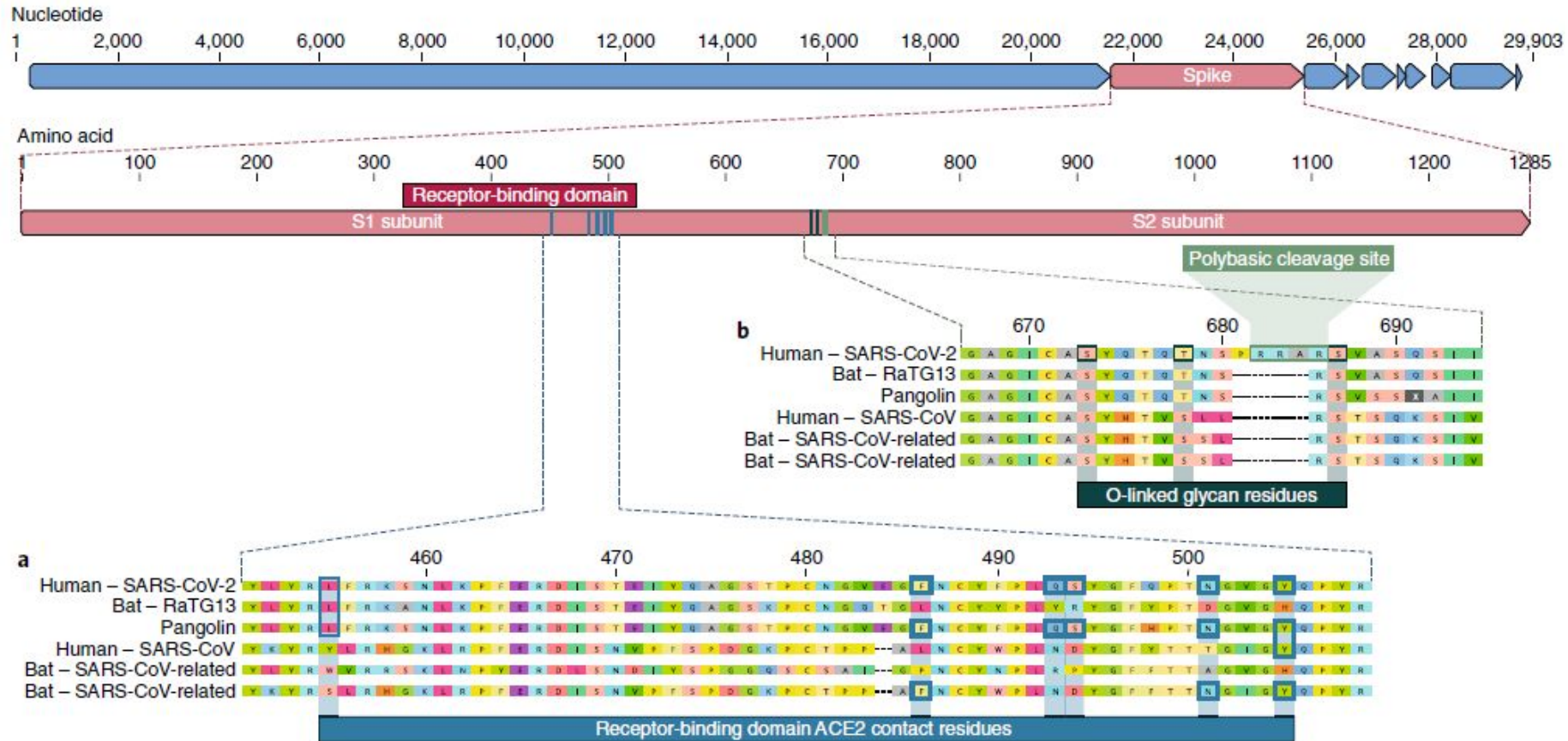
Lifecycle



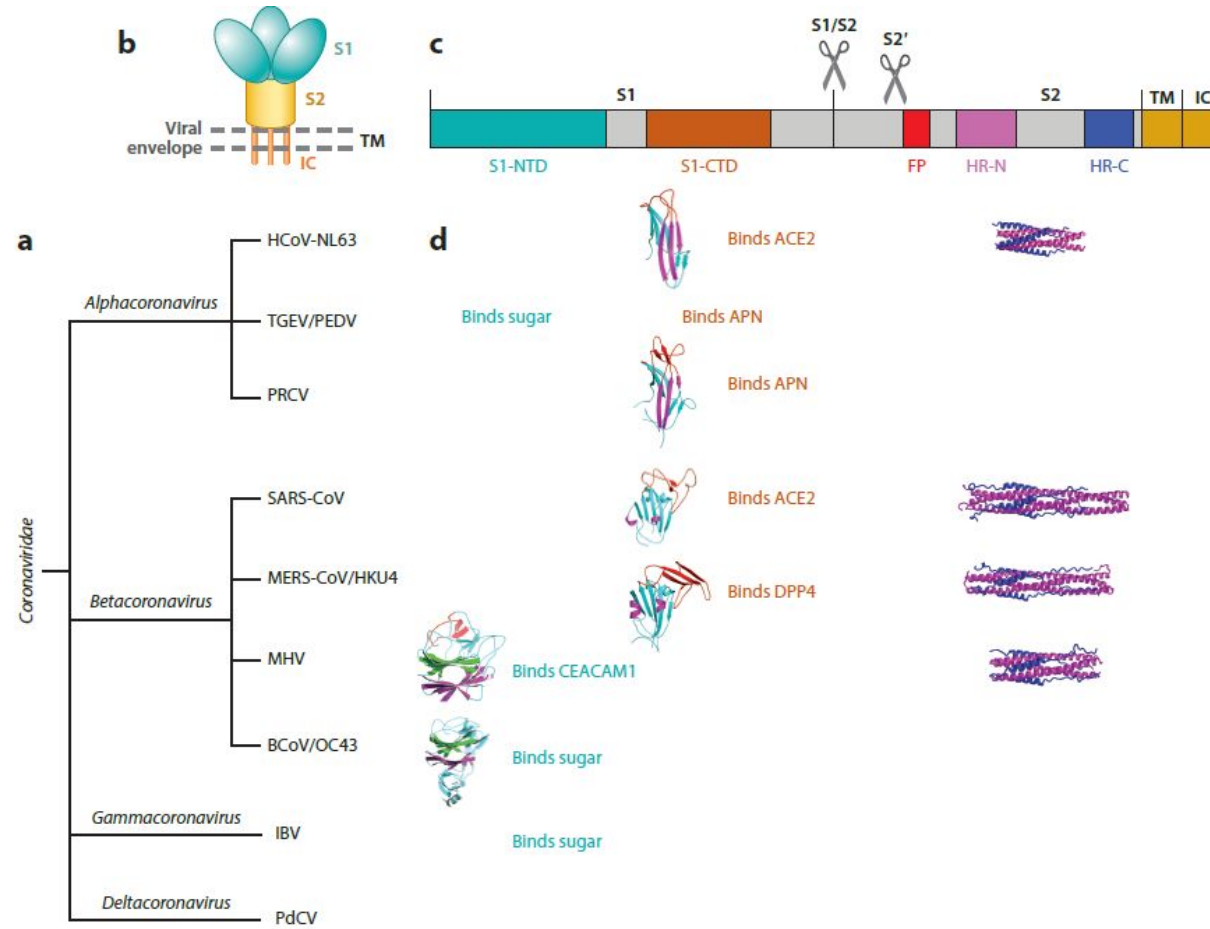
Cui, J., Li, F. & Shi, Z. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181–192 (2019).
<https://doi.org/10.1038/s41579-018-0118-9>



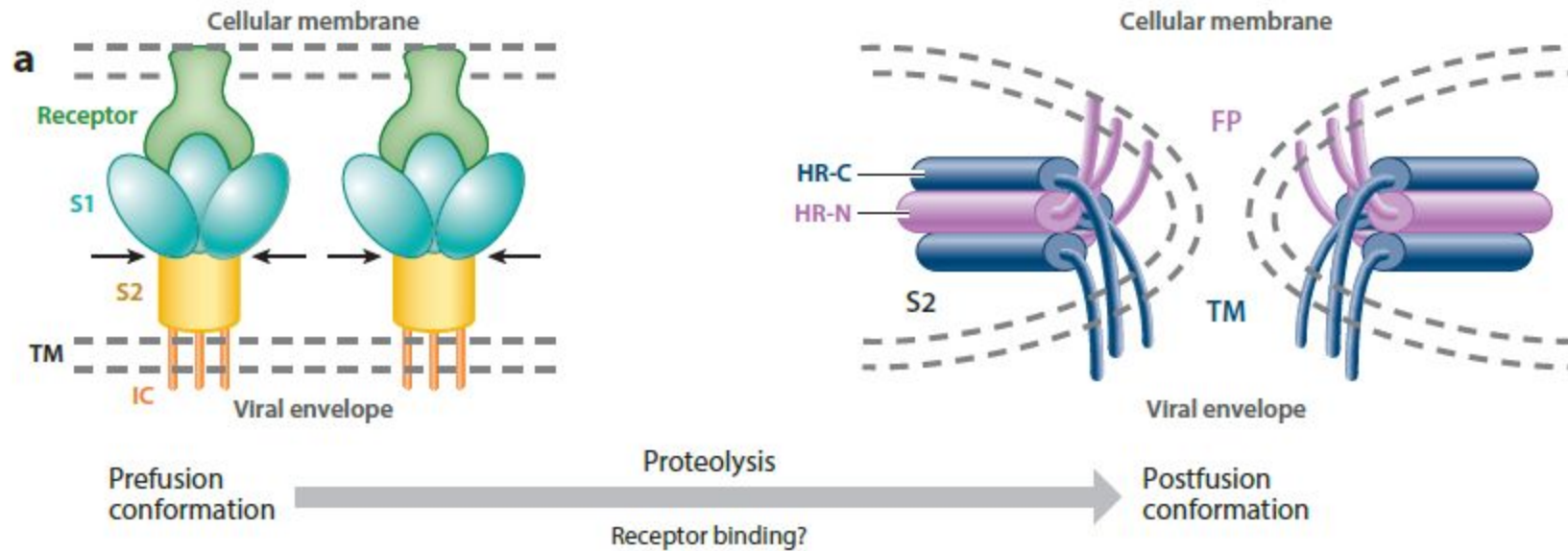
Coronavirus Genome Structure



S (spike) Protein Diversity and Similarity



Membrane Fusion



Why is all this important?

Just like the presence or absence of a virulence or AMR gene can significantly impact public health decision making, so can a single nucleotide change in a virus.

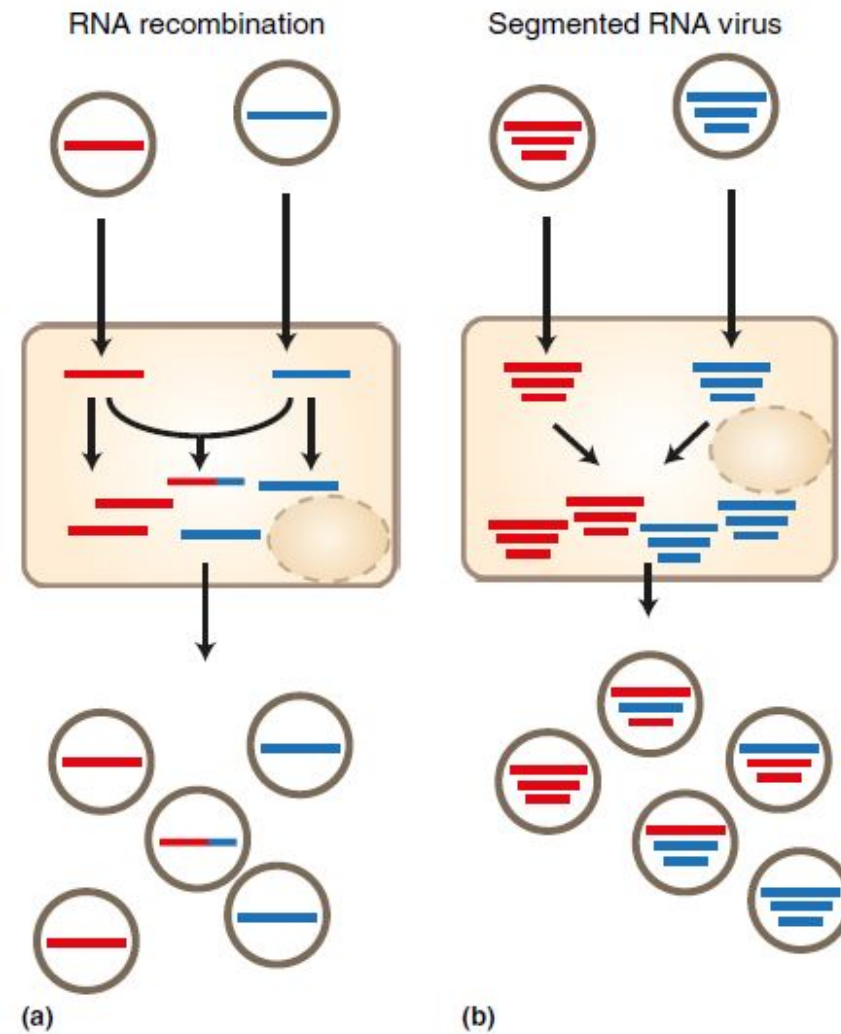
For SARS-CoV-2, it is not just the phylogeny which is important, but each individual change itself should be examined.

Not a lot of gene content changes.

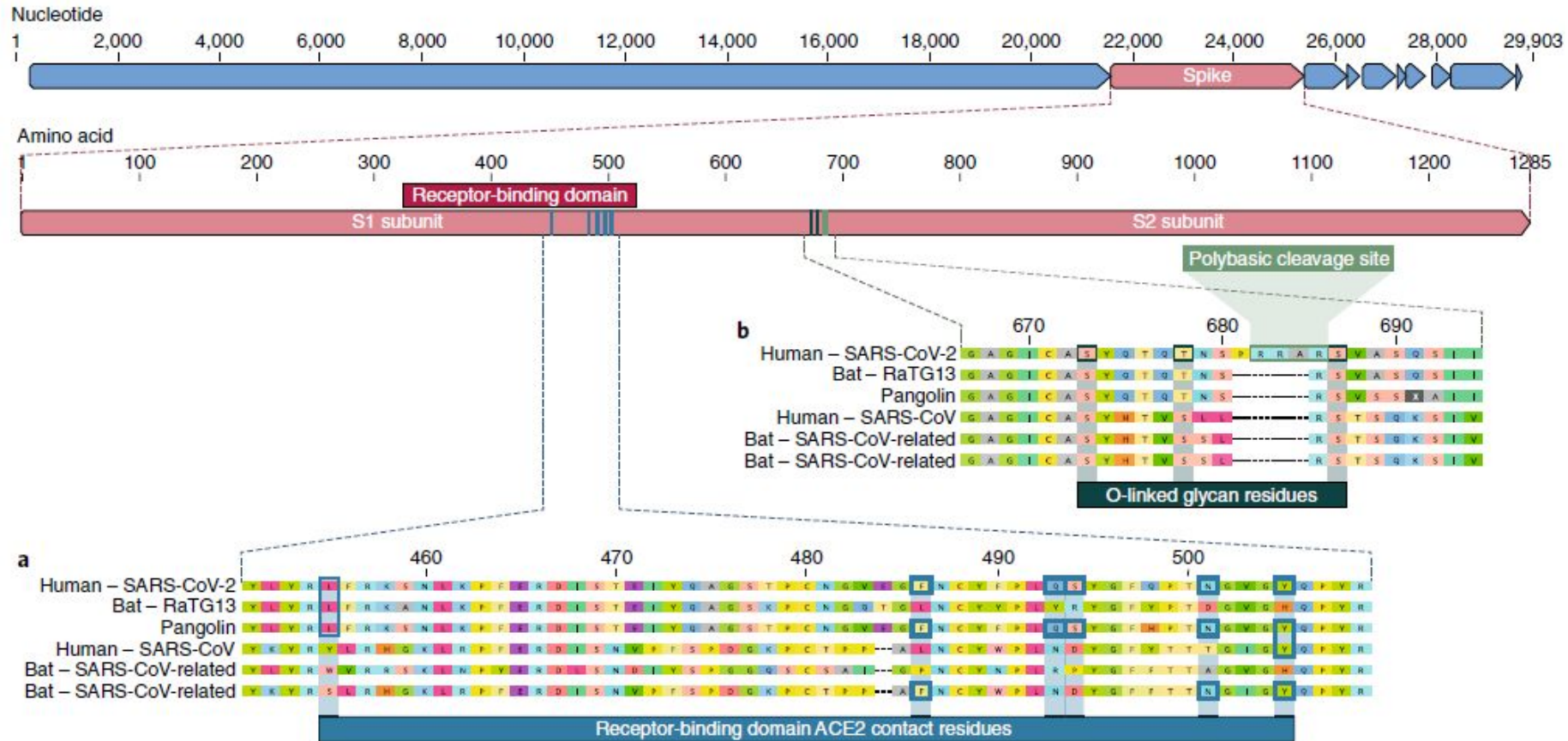


Recombination

Figure 5 RNA recombination and reassortment in RNA viruses. (a) RNA recombination. This occurs following co-infection of a cell by different virus strains and can lead to the generation of recombinant viruses through copy-choice replication. RNA recombination can occur in both unsegmented viruses (as shown) or within the segment of a segmented virus. (b) Reassortment. This occurs following co-infection of a cell by different strains of a segmented virus, and can generate different combinations of reassortant progeny. Adapted from Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? *Nature Reviews Microbiology* 9, 617–626.

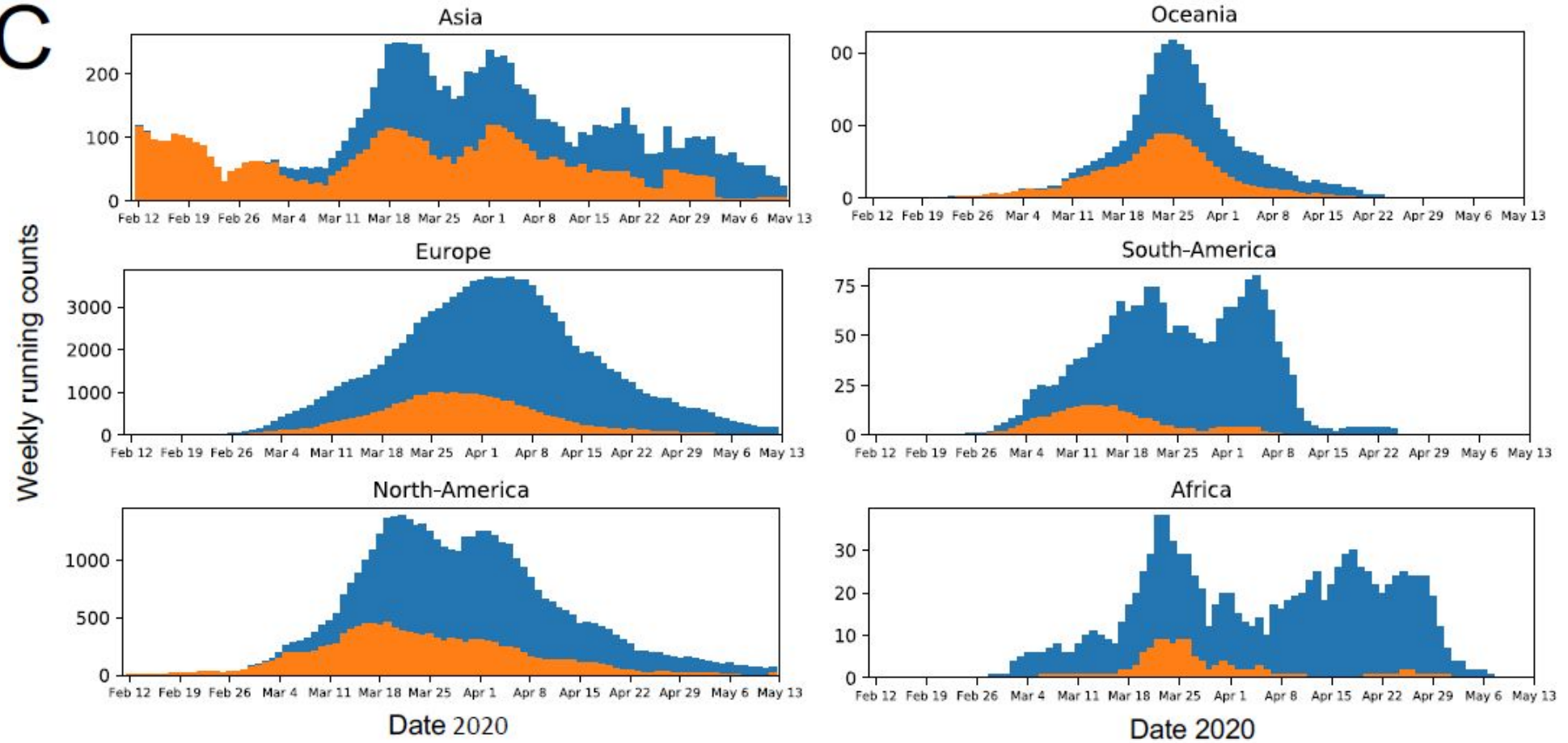


Coronavirus Genome Structure



D614G Variant

C



Discussion

