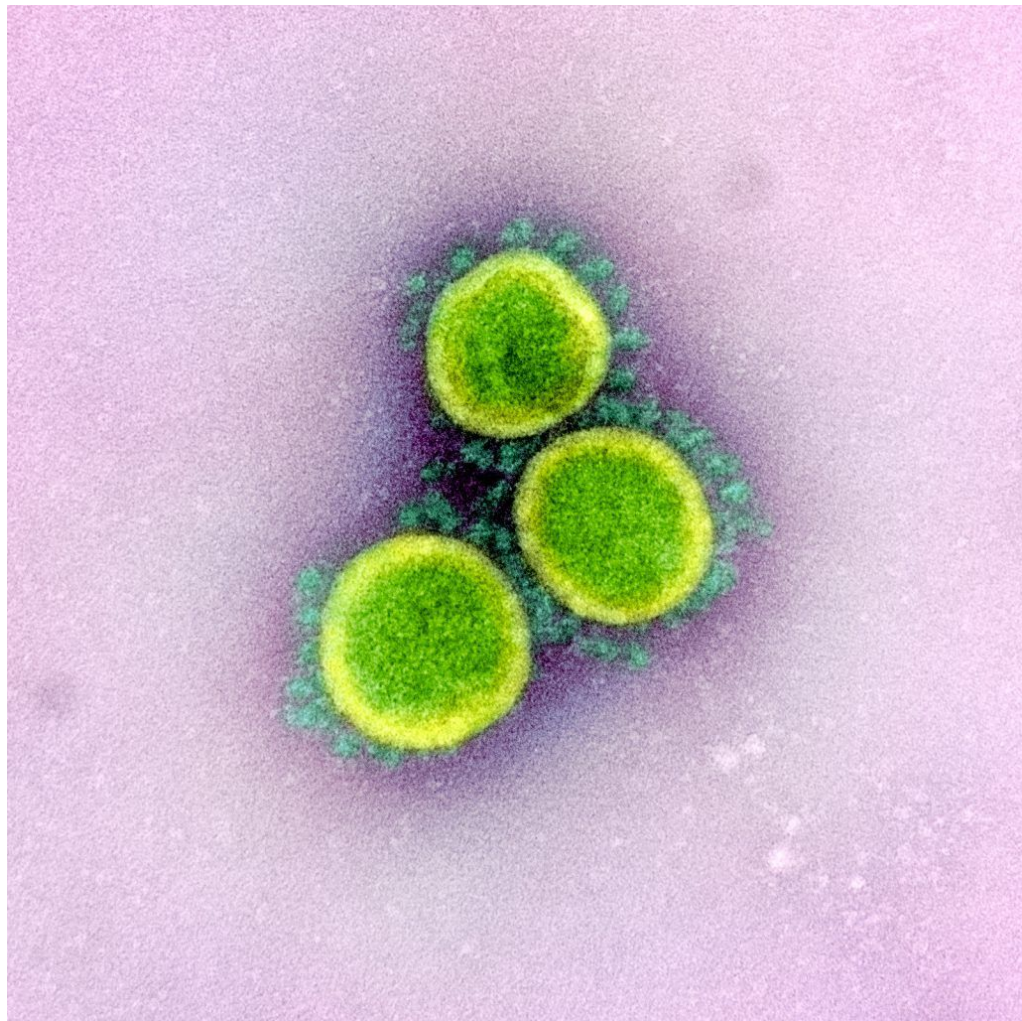
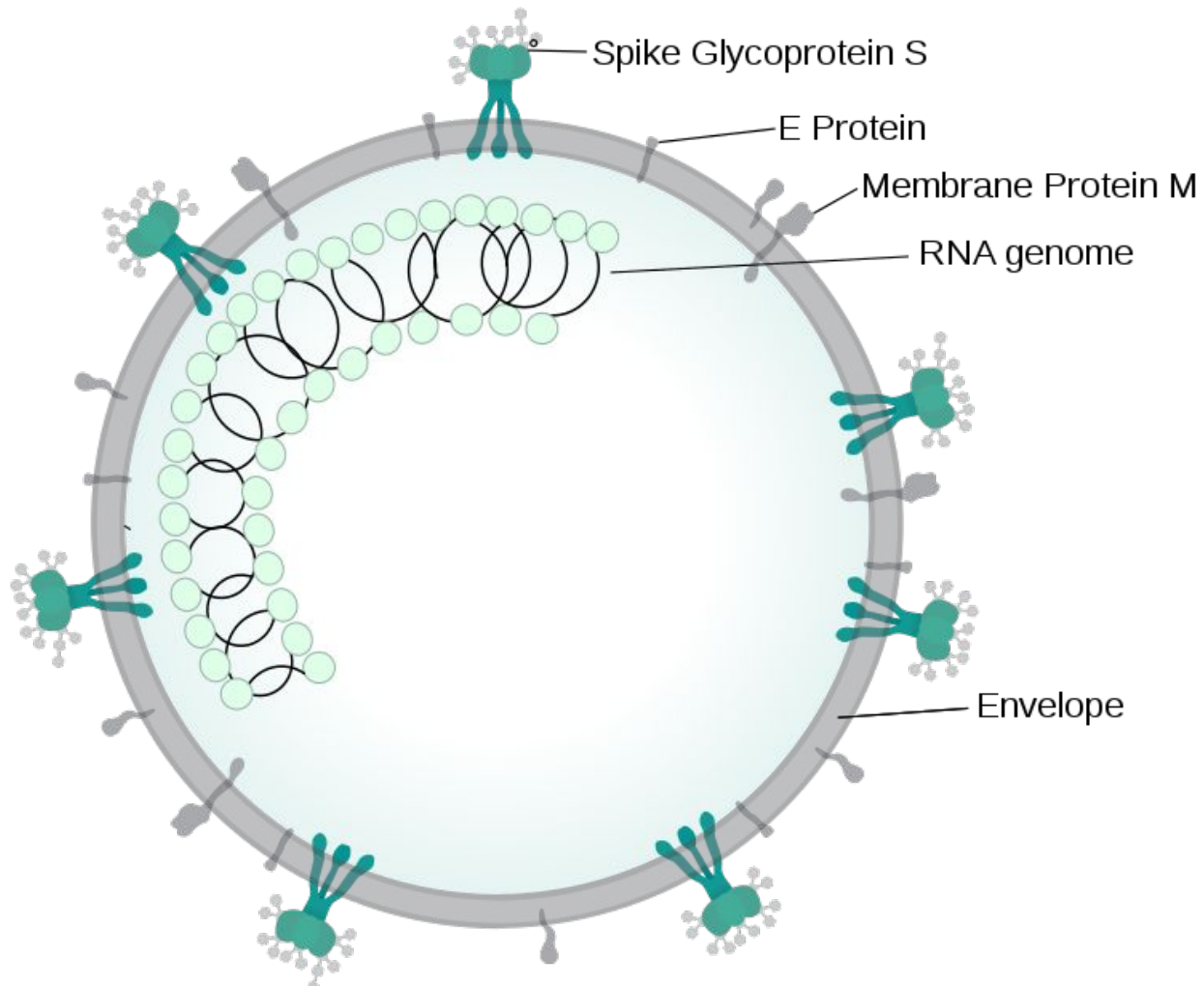


A SMALL PEEK INTO SARS-COV-2 GENOMICS

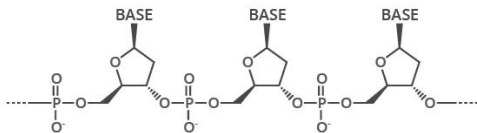
Melih Yıldız
Middle East Technical University





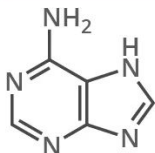
THE CHEMICAL STRUCTURE OF DNA

THE SUGAR PHOSPHATE 'BACKBONE'

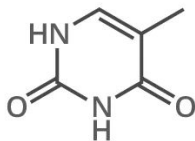


DNA is a polymer made up of units called nucleotides. The nucleotides are made of three different components: a sugar group, a phosphate group, and a base. There are four different bases: adenine, thymine, guanine and cytosine.

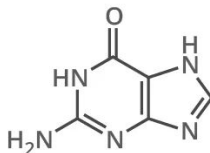
A ADENINE



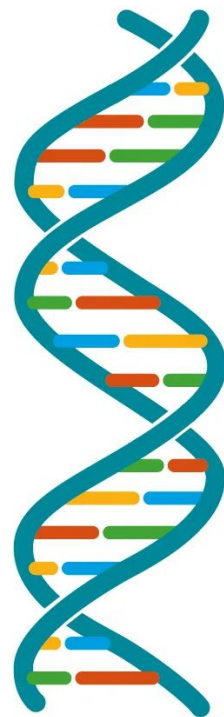
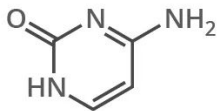
T THYMINE



G GUANINE

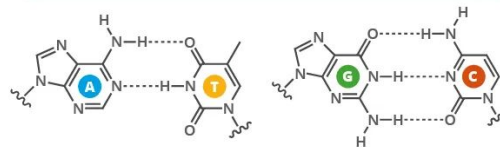


C CYTOSINE



WHAT HOLDS DNA STRANDS TOGETHER?

DNA strands are held together by hydrogen bonds between bases on adjacent strands. Adenine (A) always pairs with thymine (T), while guanine (G) always pairs with cytosine (C). Adenine pairs with uracil (U) in RNA.



FROM DNA TO PROTEINS

The bases on a single strand of DNA act as a code. The letters form three letter codons, which code for amino acids - the building blocks of proteins.



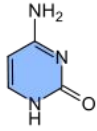
An enzyme, RNA polymerase, transcribes DNA into mRNA (messenger ribonucleic acid). It splits apart the two strands that form the double helix, then reads a strand and copies the sequence of nucleotides. The only difference between the RNA and the original DNA is that in the place of thymine (T), another base with a similar structure is used: uracil (U).



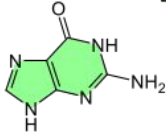
In multicellular organisms, the mRNA carries genetic code out of the cell nucleus, to the cytoplasm. Here, protein synthesis takes place. 'Translation' is the process of turning the mRNA's 'code' into proteins. Molecules called ribosomes carry out this process, building up proteins from the amino acids coded for.



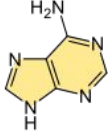
CYTOSINE **C**



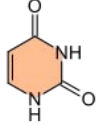
GUANINE **G**



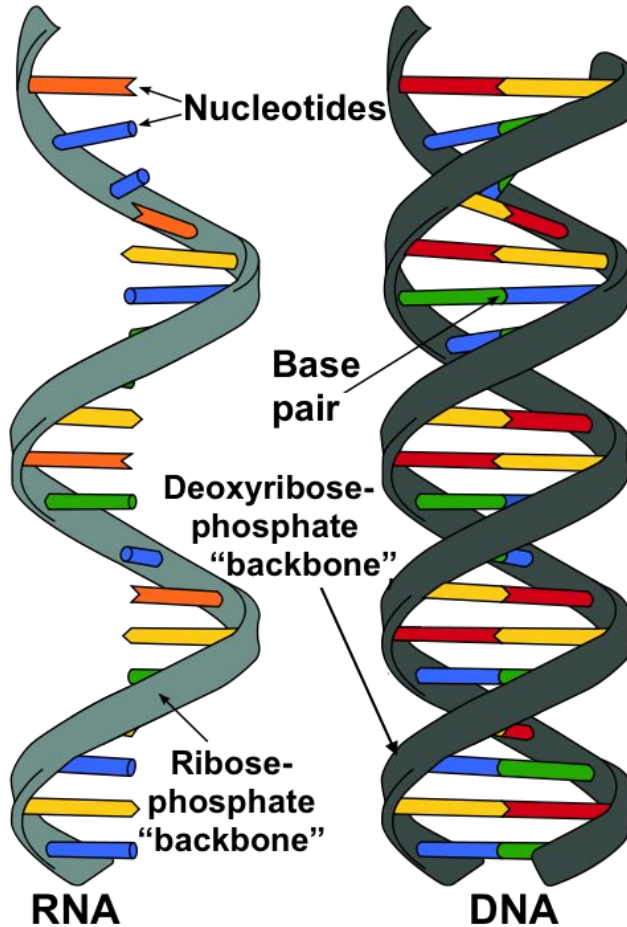
ADENINE **A**



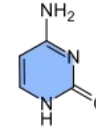
URACIL **U**



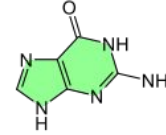
**Nucleotides
of RNA**



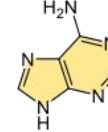
CYTOSINE **C**



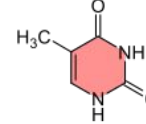
GUANINE **G**



ADENINE **A**



THYMINE **T**



**Nucleotides
of DNA**

(ribonucleic acid) (deoxyribonucleic acid)



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: <https://www.coronavirus.gov>.
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.



Severe acute respiratory syndrome coronavirus 2 data hub

Search, retrieve, and analyze SARS-CoV-2 GenBank data.

- [Betacoronavirus BLAST](#)
- [SARS-CoV-2 articles in PubMed](#)
- [CDC outbreak information](#)

Refine Results

Reset

Virus



Severe acute respiratory syndrome coronavirus 2, taxid:2697049

Selected Results: 0

PubMed

Download

Align

Build Phylogenetic Tree

Expand Table	Nucleotide (932)		Protein (9,894)			Select Columns	
	<input type="checkbox"/>	Accession ▾	Release Date ▾	Species ▾	Length ▾	Geo Location ▾	Host
<input type="checkbox"/>	NC_045512	2020-01-13	Severe acute respiratory s...	29903	China		

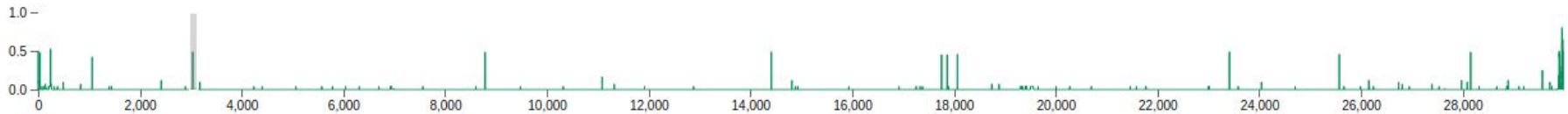
Feedback

668 complete genome sequences from 21 countries

including one genome sequence from Turkey

ATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTT
CAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAAATGGTGTGAAGGTTTTAATTGTT
ACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATAACA
GAGTAGTAGTACTTTCTTTTGAACCTTACATGCACCAGCAACTGTTTGTGGACCTAAAA
AGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTCAATGGTTTAAACAGGCA
CAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCACAATTTGGCAGAGACA
TTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACAC
CATGTTCTTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAACCAGGTTG
CTGTTCTTTATCAGGGTGTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAAC
TTACTCCTACTTGGCGTGTATTTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCT
GTTTAAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGACATACCCATTGGTGCAG
GTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTA
GTCAATCCATCATTGCCTACACTATGCACTTGGTGCAGAAAATTCAGTTGCTTACTCTA
ATAACTCTATTGCCATACCCACAAAATTTACTATTAGTGTACCACAGAAAATTCACCAG
TGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAAT
GCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCCTGCTTTAACTG
GAATAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTTGCACAAGTCAAACAAATTT
ACAAAACACCACCAATTAAGATTTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATC
CATCAAACCAAGCAAGAGGTCAATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTG
CAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACC
TCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAA
TGATTGCTCAATACACTTCTGCACTGTAGCGGGTACAATCACTTCTGTTTGGACCTTTG
GTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAAATGGTA
TTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATGATTGCCAACCAATTTAATA
GTGCTATTGGCAAATTAAGACTCACTTTCTTCCACAGCAAGTGAAGTGGAAAATTC
AAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACCTTAGCTCCA
ATTTTGGTGAATTTCAAGTGTTTAAATGATATCCTTTCACGTCTTGACAAAGTTGAGG
CTGAAGTGCAATTTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGA
CTCAACAATTAATTAGAGCTGCAGAAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAAA
TGTCAGAGTGTGACTTGGACAATCAAAAAGAGTTGATTTTTTGTGGAAAGGGCTATCATC
TATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGATGTGACTTATGTC
CTGCACAAGAAAAGAACTTCACAACCTGCTCCTGCCATTTGTATGATGGAAAAGCACACT
TTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAAACAAAAGGAATT

SEQUENCE OVERVIEW



ORF1a

ORF1b

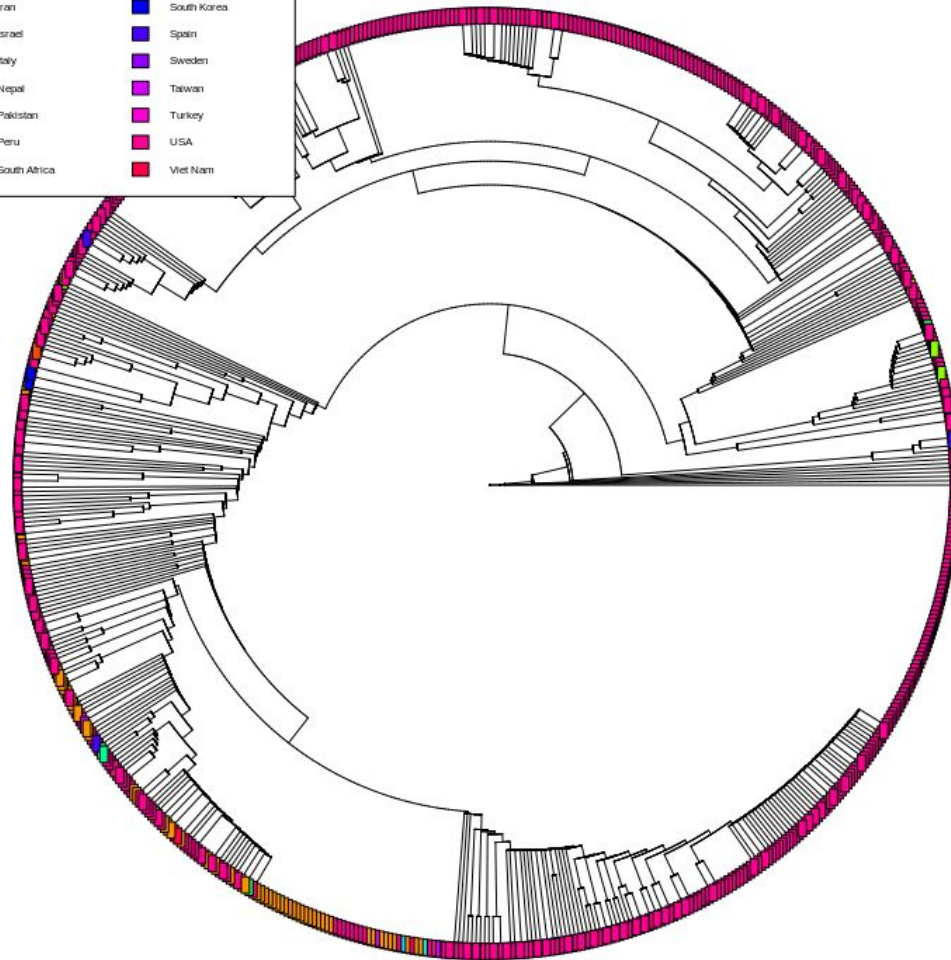
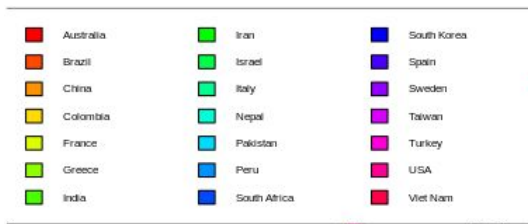
S

Total of 10 genes, some are marked

Phylogenetic Relationships of Virus Isolates

Two phylogenetic tree methods:

- UPGMA
- Neighbor-Joining



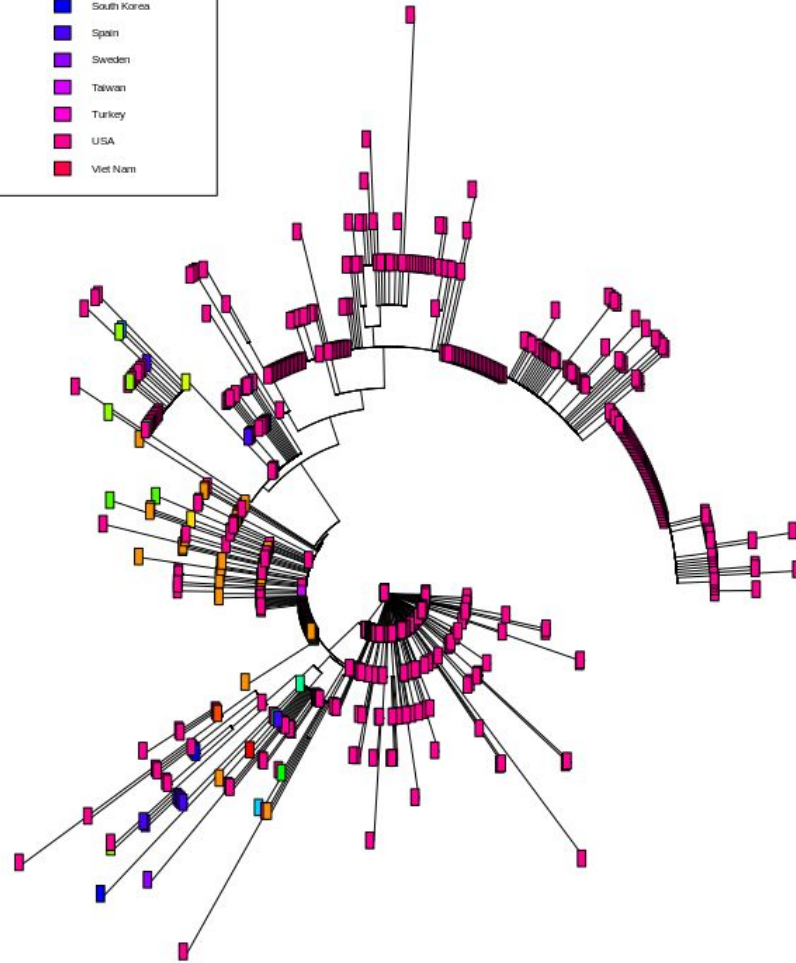
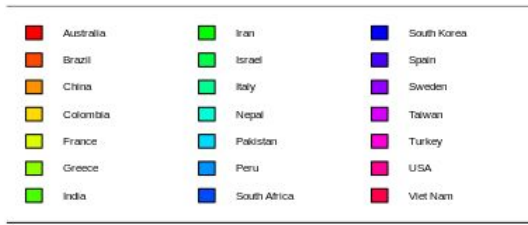
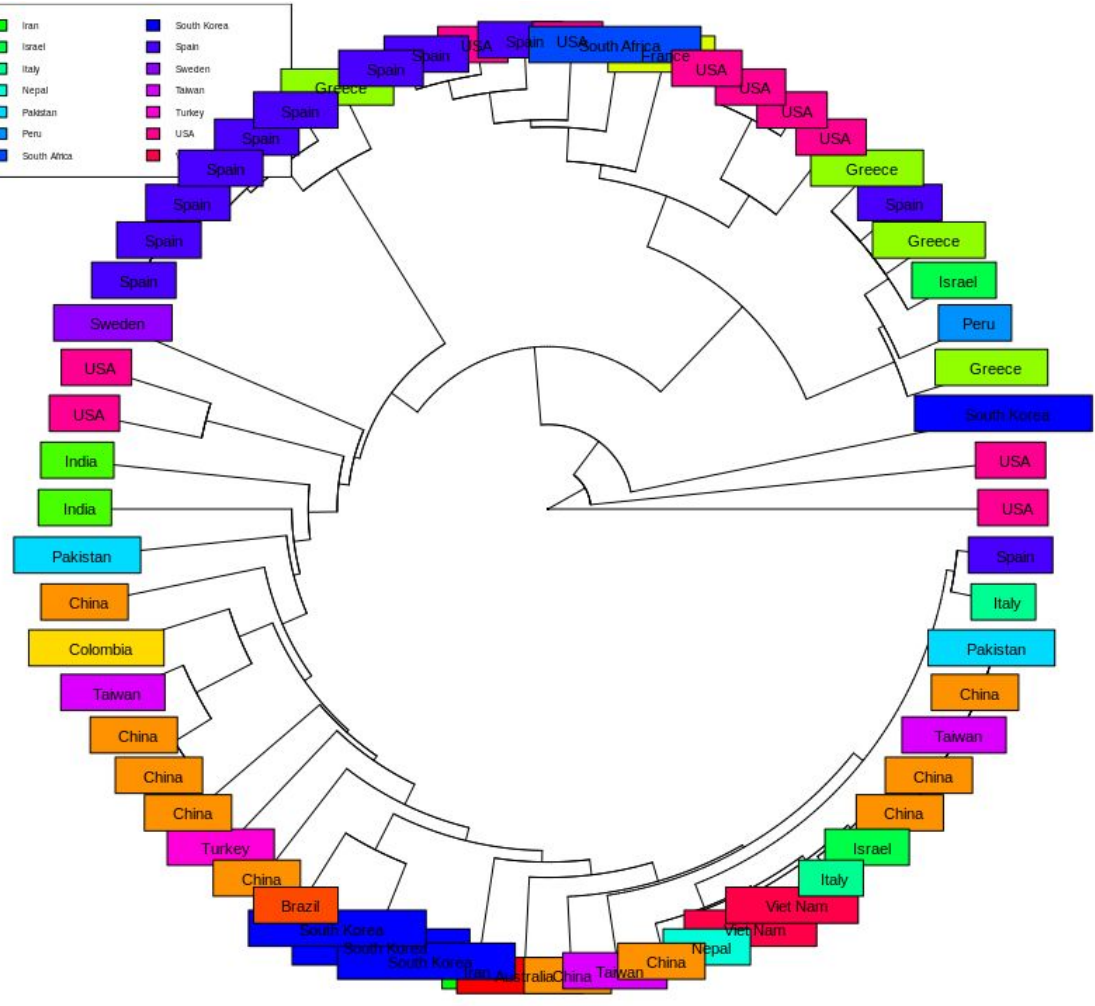
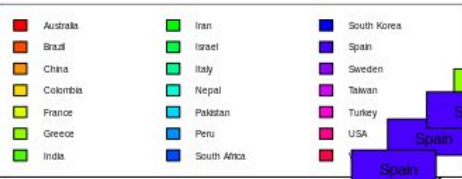
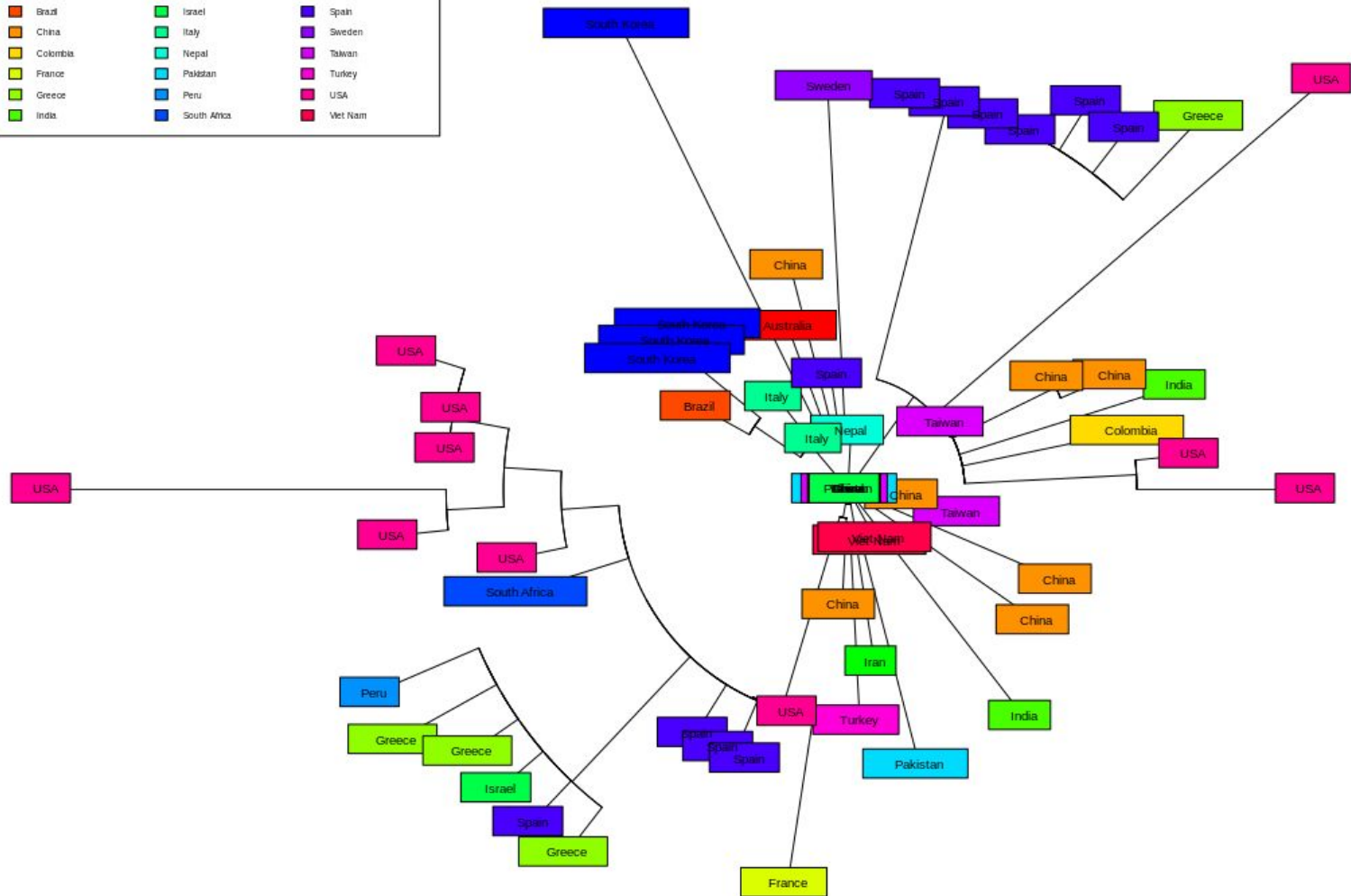
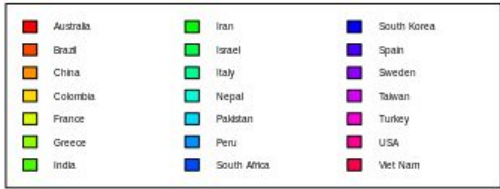


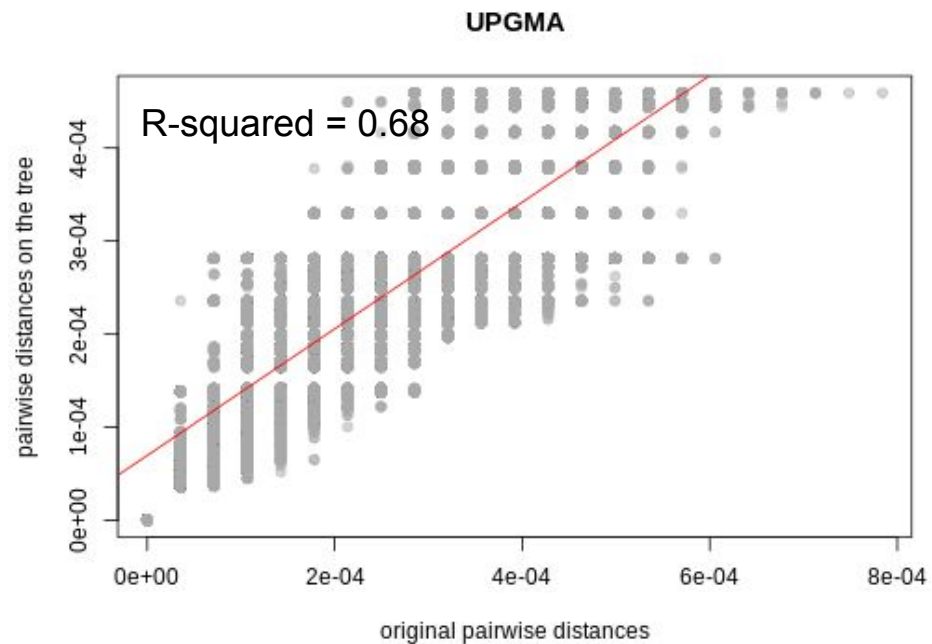
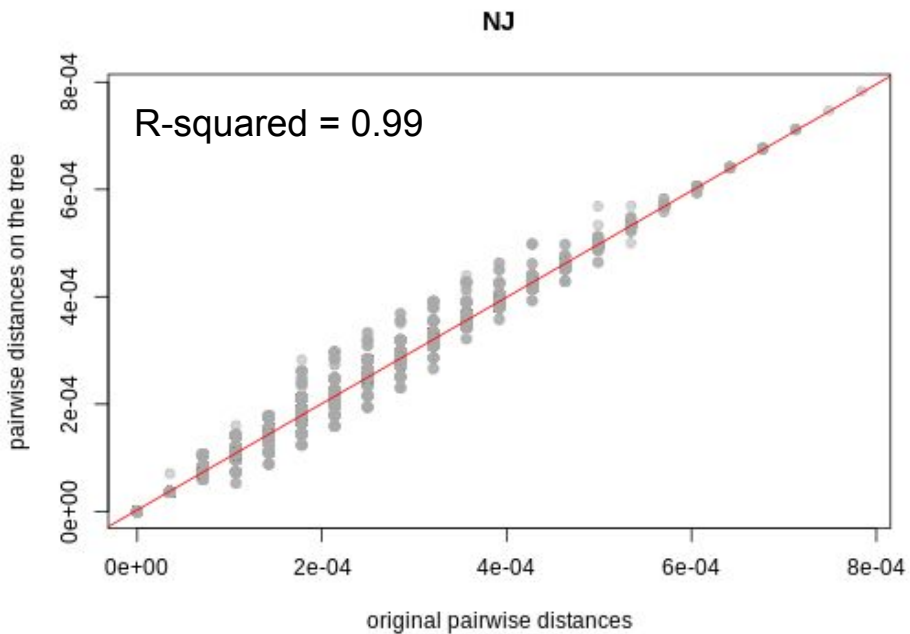
Table 1: Number of full genome sequences used in the analysis.

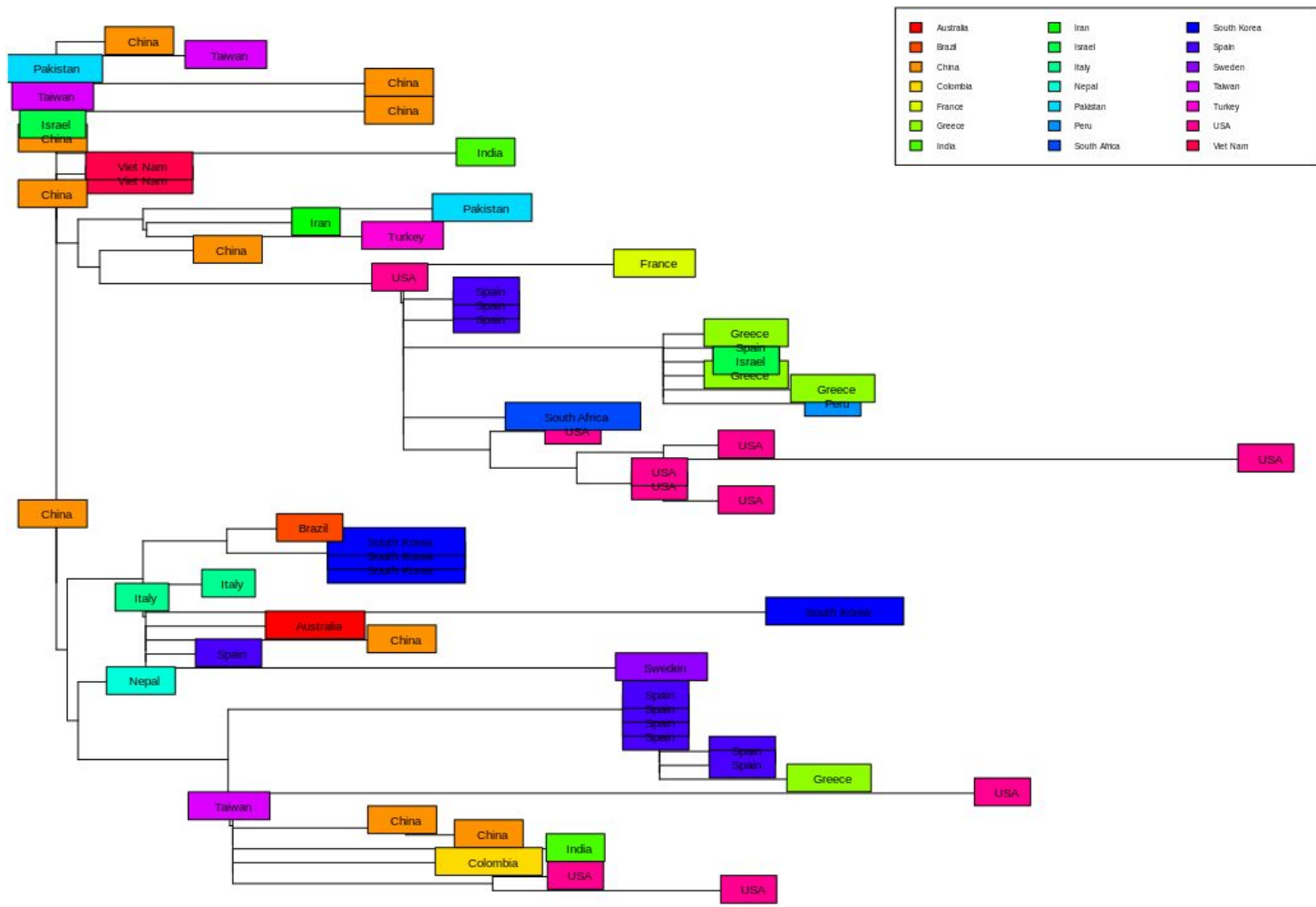
Country of Sequence Origin	Number of Sequences	Country of Sequence Origin	Number of Sequences
USA	564	Australia	1
China	61	Brazil	1
Spain	11	Colombia	1
Greece	4	France	1
South Korea	4	Iran	1
Taiwan	3	Nepal	1
India	2	Peru	1
Isreal	2	South Africa	1
Italy	2	Sweden	1
Pakistan	2	Turkey	1
Vietnam	2		





WHICH TREE IS BETTER?





Further Directions

- Observing evolution of the virus in real time
- Identification of present virus subtypes in different regions
*Important for qPCR, diagnostics and possibly vaccine development
- Analysis of conservation, drift and positive selection on viral genes.

Thank you!

Reference

- NCBI Virus Database(<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>)
- Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., ... Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), 265–269. doi: 10.1038/s41586-020-2008-3
- <https://www.khanacademy.org/science/high-school-biology/hs-molecular-genetics/hs-rna-and-protein-synthesis/a/hs-rna-and-protein-synthesis-review>