



Mutation at the 501 Position of SARS-CoV-2 Results in a Higher Binding Affinity to the Human ACE2 Receptor

Kam Taghizadeh, Fatima Alghanem



Outline

- Interaction of SARS-CoV-2 and SARS-CoV to human ACE2
- Future evolution of SARS-CoV-2
- Mutation at position 501 leads to a higher infectious strain
- Our research question: Why does the viral strain become more infectious if there is a mutation at position 501?
- Our methods and results of creating a sequence alignment and phylogenetic tree
- Future directions we would take

Interaction of SARS-CoV-2 and ACE2

- SARS-CoV-2 spike protein receptor-binding domain (RBD) specifically recognizes its host receptor angiotensin-converting enzyme 2 (ACE2) (Wan et al, 2020)
- Research has shown that the specific host that gets infected is dependent on the binding affinity between the viral RBD and host ACE2 in the primary viral attachment step (Wan et al, 2020)

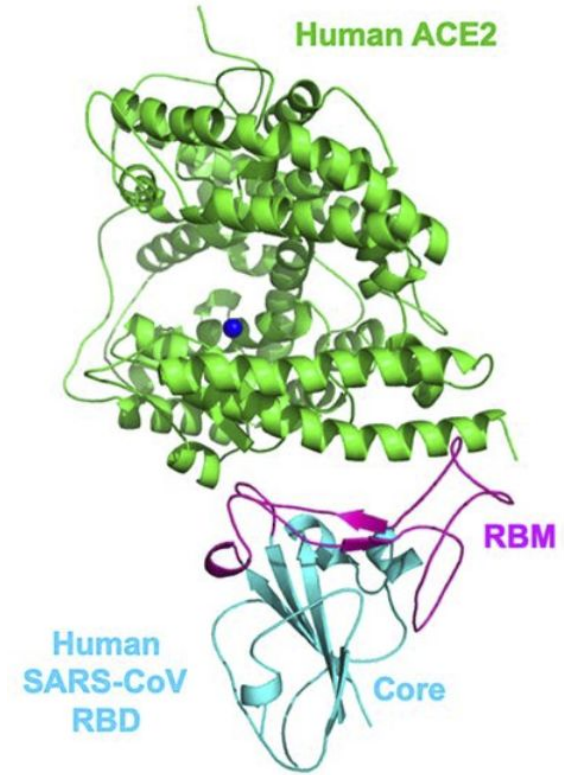


Figure 1: Human SARS-CoV: Human ACE2 (green), RBM (magenta), core (blue)

Interaction of SARS-CoV-2 and ACE2

- Two virus-binding hotspots have been discovered on human ACE2
 - Various mutations have occurred near these hotspots, including residue 487 in SARS-CoV, which corresponds to residue 501 in SARS-CoV-2 (Wan et al, 2020)
 - Residue 487 in SARS-CoV (501 in SARS-CoV-2) is known to enhance binding affinity to human ACE2 (Wan et al, 2020)
 - Mutation of residue 487 in SARS-CoV strengthens the structural stability and enhances the viral binding to human Hotspot Lys353 in ACE2 (Wan et al, 2020)

Virus	Year	442	472	479	480	487
SARS - human	2002	Y	L	N	D	T
SARS - civet	2002	Y	L	K	D	S
SARS - human/civet	2003	Y	P	N	G	S
SARS - civet	2005	Y	P	R	G	S
SARS - human	2008	F	F	N	D	S
Viral adaption to human ACE2		F > Y	F > L > P	N = R >>> K	D > G	T >>> S
Optimized - human	In vitro design	F	F	N	D	T
Viral adaptation to civet ACE2		Y > F	P = L > F	R > K = N	G > D	T > S
Optimized - civet	In vitro design	Y	P	R	G	T
SARS - bat	2013	S	F	N	D	N
2019-nCoV – human	2019	L (455)	F (486)	Q (493)	S (494)	N (501)

Figure 2: Comparison of the viral strains, note position 487 and its correspondence to position 501 in 2019-nCoV-human



Outline

- Interaction of SARS-CoV-2 and SARS-CoV to human ACE2
- Mutation at position 501 leads to a higher infectious strain
- Our research question: Why does the viral strain become more infectious if there is a mutation at position 501?
- Our methods and results of creating a sequence alignment and phylogenetic tree
- Future directions we would take

Future Evolution of SARS-CoV-2

- Asn501 is among the top twenty residues that will most likely mutate and have an adversarial effect on infectivity

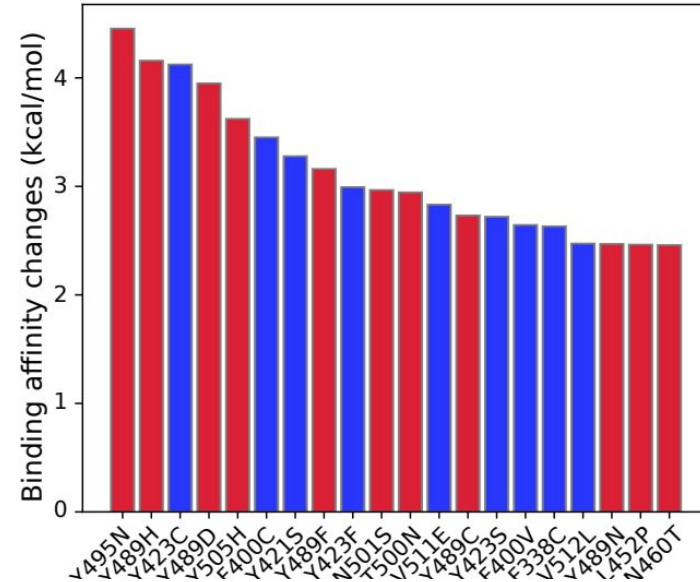
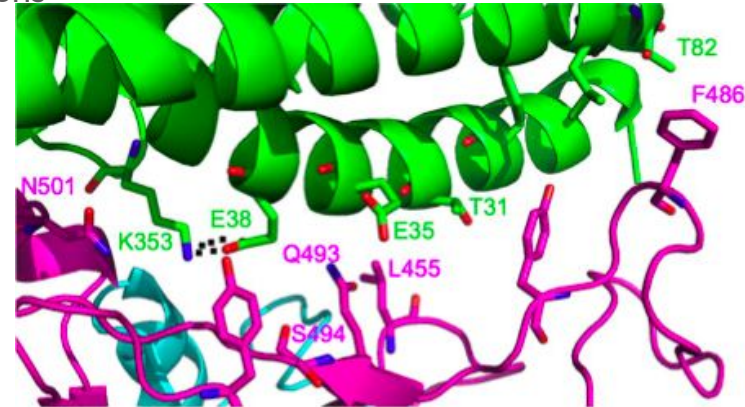


Figure 3: Top 20 amino acids most likely to mutate (Red=mutations on RBD, blue=mutations away from RBD) <https://arxiv.org/pdf/2005.14669.pdf>

Research Question- Why does the viral strain become more infectious if there is a mutation at position 501?

- According to Chunyan Yi et al. , the enhanced binding affinity may be due to greater support provided by residue 501 in SARS-CoV-2 or 487 in SARS-CoV for the stabilization of the overall structure of RBD, or the strengthening of hydrophilic reactions





We Wanted to Find Strains that Have Evolved at Position 501

- We decided to see if we could find a SARS-CoV-2 spike protein sequence that would have a mutation at position 501
- We visited the NCBI Virus SARS-CoV-2 Data Hub to collect two sequences each from different geographic regions
 - United Kingdom, Spain, Iran, Saudi Arabia, Italy, Venezuela, Peru, and Brazil
- All of the sequences were downloaded in FASTA format
 - We used google character count to see what the nucleotide for each sequence was at position 501
- We then created a clustal alignment by visiting phylogeny.fr
 - We scrolled down the page to the section called 'Phylogeny analysis', and clicked on the text 'One Click'.
 - Uploaded the set of sequences in FASTA format into the large text box, and clicked submit
 - Clicked on '3. Alignment', then under 'outputs' clicked on 'Alignment in Clustal format'
 - Results are shown on next few slides

QIM47476.1 MFVFLVLLPLVSSQCENLTTRTQLPPAYTNSITRFGVYYPDKVFRSSVLHSTQDLFLPFFS
 QOJ75919.1 MFVFLVLLPLISSQCENLTTRTQLPPAYTNSITRFGVYYPDKVFRSSVLHSTQDLFLPFFS
 QMX86941.1 MFVFLVLLPLVSSQCENLTTRTQLPPAYTNSITRFGVYYPDKVFRSSVLHSTQDLFLPFFS
 QMX86953.1 MFVFLVLLPLVSSQCENLTTRTQLPPAYTNSITRFGVYYPDKVFRSSVLHSTQDLFLPFFS
 QIM47467.2 MFVFLVLLPLVSSQCENLTTRTQLPPAYTNSITRFGVYYPDKVFRSSVLHSTQDLFLPFFS

QIM47476.1 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIIV
 QOJ75919.1 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIIV
 QMX86941.1 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIIV
 QMX86953.1 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIIV
 QIM47467.2 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIIV

QIM47476.1 NNATNVVIVKCEQFCNDPFLGVYYHKNKSMWSEFRVSSANNCTFEYVSQPFIMDLME
 QOJ75919.1 NNATNVVIVKCEQFCNDPFLGVYYHKNKSMWSEFRVSSANNCTFEYVSQPFIMDLME
 QMX86941.1 NNATNVVIVKCEQFCNDPFLGVYYHKNKSMWSEFRVSSANNCTFEYVSQPFIMDLME
 QMX86953.1 NNATNVVIVKCEQFCNDPFLGVYYHKNKSMWSEFRVSSANNCTFEYVSQPFIMDLME
 QIM47467.2 NNATNVVIVKCEQFCNDPFLGVYYHKNKSMWSEFRVSSANNCTFEYVSQPFIMDLME

QIM47476.1 GKQGNFKNLRREFVFNKIDGYFKIYSKHTPINLVRDLPGQFSALEPLVDLPIGINITRFQT
 QOJ75919.1 GKQGNFKNLRREFVFNKIDGYFKIYSKHTPINLVRDLPGQFSALEPLVDLPIGINITRFQT
 QMX86941.1 GKQGNFKNLRREFVFNKIDGYFKIYSKHTPINLVRDLPGQFSALEPLVDLPIGINITRFQT
 QMX86953.1 GKQGNFKNLRREFVFNKIDGYFKIYSKHTPINLVRDLPGQFSALEPLVDLPIGINITRFQT
 QIM47467.2 GKQGNFKNLRREFVFNKIDGYFKIYSKHTPINLVRDLPGQFSALEPLVDLPIGINITRFQT

QIM47476.1 LLAHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKNENGITIDAVDCALDPLSETK
 QOJ75919.1 LLAHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKNENGITIDAVDCALDPLSETK
 QMX86941.1 LLAHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKNENGITIDAVDCALDPLSETK
 QMX86953.1 LLAHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKNENGITIDAVDCALDPLSETK
 QIM47467.2 LLAHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKNENGITIDAVDCALDPLSETK

QIM47476.1 XXXXXXXXXXXXXXXXXXXXXXXXRFNITNLCPFGEVFNATRFASVYAWNKRKISN
 QOJ75919.1 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKISN
 QMX86941.1 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKISN
 QMX86953.1 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKISN
 QIM47467.2 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKISN

QIM47476.1 CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD
 QOJ75919.1 CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD
 QMX86941.1 CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD
 QMX86953.1 CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD
 QIM47467.2 CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD

QIM47476.1 YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFERDISTEIIYQAGSTPC
 QOJ75919.1 YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFERDISTEIIYQAGSTPC
 QMX86941.1 YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFERDISTEIIYQAGSTPC
 QMX86953.1 YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFERDISTEIIYQAGSTPC
 QIM47467.2 YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFERDISTEIIYQAGSTPC

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

QIM47476.1
 QOJ75919.1
 QMX86941.1
 QMX86953.1
 QIM47467.2

NGVEGFNCYFPLQSYGFQPTNGVGYPYRVVLSFELLHAPATVCGPKKSNLVKNKCVN
 NGVEGFNCYFPLQSYGFYTTNGVGYPYRVVLSFELLHAPATVCGPKKSNLVKNKCVN
 NGVEGFNCYFPLQSYGFQPTNGVGYPYRVVLSFELLHAPATVCGPKKSNLVKNKCVN
 NGVEGFNCYFPLQSYGFQPTNGVGYPYRVVLSFELLHAPATVCGPKKSNLVKNKCVN
 NGVEGFNCYFPLQSYGFQPTNGVGYPYRVVLSFELLHAPATVCGPKKSNLVKNKCVN

FNFNGLTGTGVLTESNKKFLPFQFGRIADTTDAVRDPQTLEILDIIPCFSFGGVSVITP
 FNFNGLTGTGVLTESNKKFLPFQFGRIADTTDAVRDPQTLEILDIIPCFSFGGVSVITP
 FNFNGLTGTGVLTESNKKFLPFQFGRIADTTDAVRDPQTLEILDIIPCFSFGGVSVITP
 FNFNGLTGTGVLTESNKKFLPFQFGRIADTTDAVRDPQTLEILDIIPCFSFGGVSVITP
 FNFNGLTGTGVLTESNKKFLPFQFGRIADTTDAVRDPQTLEILDIIPCFSFGGVSVITP

GTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAHEVNNSY
 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAHEVNNSY
 GTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAHEVNNSY
 GTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAHEVNNSY
 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAHEVNNSY

ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPNTFTI
 ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPNTFTI
 ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPNTFTI
 ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPNTFTI
 ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNSIAIPNTFTI

- “*” for invariant
- “.” for highly conserved
- “.” for weakly conserved
- "space" for not conserved

QIM47476.1	SVTTEILPVSMTKTSVDCTMYICGDSSTECSNLLLOYGSFCTQLNRALTGIAVEQDKNTQE
QOJ75919.1	SVTTEFLPVSMTKTSVDCTMYICGDSSTECSNLLLOYGSFCTQLNRALTGIAVEQDKNTQE
QMX86941.1	SVTTEILPVSMTKTSVDCTMYICGDSSTECSNLLLOYGSFCTQLNRALTGIAVEQDKNTQE
QMX86953.1	SVTTEILPVSMTKTSVDCTMYICGDSSTECSNLLLOYGSFCTQLNRALTGIAVEQDKNTQE
QIM47467.2	SVTTEILPVSMTKTSVDCTMYICGDSSTECSNLLLOYGSFCTQLNRALTGIAVEQDKNTQE
	*****;
QIM47476.1	VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD
QOJ75919.1	VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD
QMX86941.1	VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD
QMX86953.1	VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD
QIM47467.2	VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD

QIM47476.1	LGDIAARDLICAQKFNGLTIVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAAIQIPFAM
QOJ75919.1	LGDIAARDLICAQKFNGLTIVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAAIQIPFAM
QMX86941.1	LGDIAARDLICAQKFNGLTIVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAAIQIPFAM
QMX86953.1	LGDIAARDLICAQKFNGLTIVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAAIQIPFAM
QIM47467.2	LGDIAARDLICAQKFNGLTIVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAAIQIPFAM

QIM47476.1	QMAYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDLSSTASALGKLDVVNQNAQALN
QOJ75919.1	QMAYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDLSSTASALGKLDVVNQNAQALN
QMX86941.1	QMAYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDLSSTASALGKLDVVNQNAQALN
QMX86953.1	QMAYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDLSSTASALGKLDVVNQNAQALN
QIM47467.2	QMAYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDLSSTASALGKLDVVNQNAQALN

QIM47476.1	TLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSQTYYTVOOLIRAAEIRA
QOJ75919.1	TLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSQTYYTVOOLIRAAEIRA
QMX86941.1	TLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSQTYYTVOOLIRAAEIRA
QMX86953.1	TLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSQTYYTVOOLIRAAEIRA
QIM47467.2	TLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSQTYYTVOOLIRAAEIRA

QIM47476.1	SANLAATKMSECVLGQSKRVDFCGKGHYLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA
QOJ75919.1	SANLAATKMSECVLGQSKRVDFCGKGHYLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA
QMX86941.1	SANLAATKMSECVLGQSKRVDFCGKGHYLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA
QMX86953.1	SANLAATKMSECVLGQSKRVDFCGKGHYLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA
QIM47467.2	SANLAATKMSECVLGQSKRVDFCGKGHYLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA

QIM47476.1	ICHDGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVIGIVNNTVYDP
QOJ75919.1	ICHDGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVIGIVNNTVYDP
QMX86941.1	ICHDGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVIGIVNNTVYDP
QMX86953.1	ICHDGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVIGIVNNTVYDP
QIM47467.2	ICHDGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVIGIVNNTVYDP

QIM47476.1	LQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QOJ75919.1	LQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QMX86941.1	LQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QMX86953.1	LQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QIM47467.2	LQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL

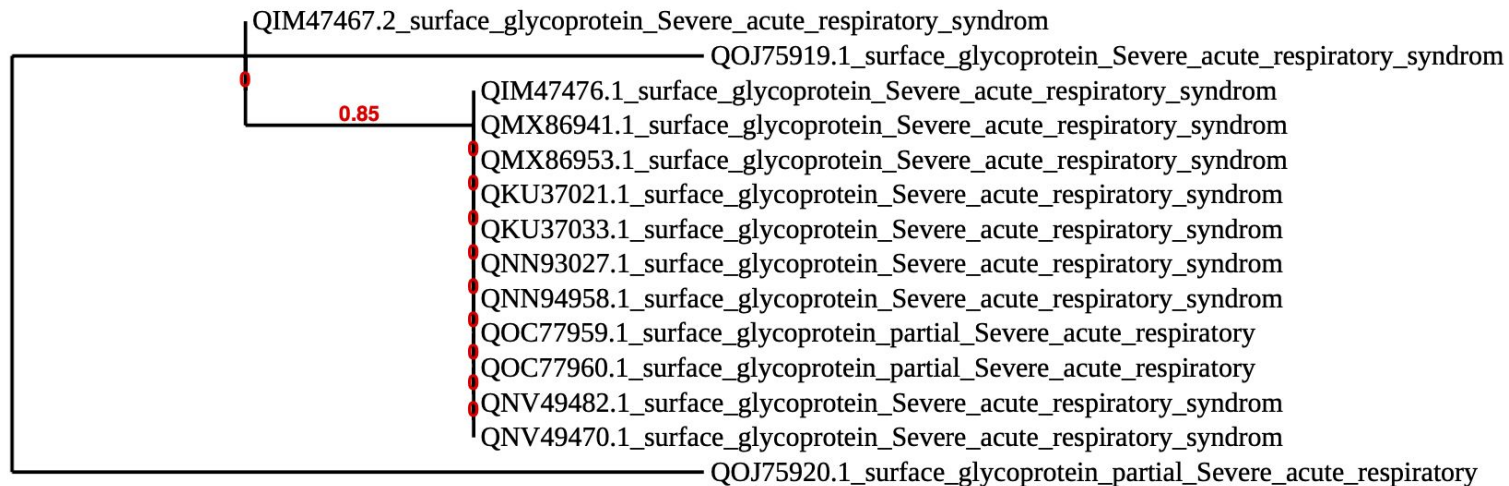
QIM47476.1	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCMTSCCSCCLKGCCSCGSCCKFDEDD
QOJ75919.1	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCMTSCCSCCLKGCCSCGSCCKFDEDD
QMX86941.1	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCMTSCCSCCLKGCCSCGSCCKFDEDD
QMX86953.1	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCMTSCCSCCLKGCCSCGSCCKFDEDD
QIM47467.2	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCMTSCCSCCLKGCCSCGSCCKFDEDD

QIM47476.1	SEPVLKGVKLHYT
QOJ75919.1	SEPVLKGVKLHYT
QMX86941.1	SEPVLKGVKLHYT
QMX86953.1	SEPVLKGVKLHYT
QIM47467.2	SEPVLKGVKLHYT



Created a Phylogenetic Tree

- We then created a phylogenetic tree by visiting phylogeny.fr
 - We scrolled down the page to the section called 'Phylogeny analysis', and clicked on the text 'One Click'.
 - Uploaded the set of sequences in FASTA format into the large text box, and clicked submit
 - Once completed, we clicked on 'tree rendering' to get the phylogenetic tree shown in the next slide



0.005



Results

- Our results showed most of the sequences having Asparagine at position 501
- The results obtained do not support our research question, due to not finding a mutation at the 501 position of any of the sequences downloaded
- Ideally the results would have shown mutations at 501 position and therefore the affinity of spike protein to Human ACE2 would have been compared between sequences that had mutation and normal sequences.



Future directions

- We look forward to researching a larger number of sequences
- Our focus would be on Sequences that have a mutation of Asparagine (n) to Threonine (t) at position 501 as previous literature showed that this kind of mutation results in a higher binding affinity to human ACE2 receptors (Wan et al, 2020)
- We want to explore the effects of other mutations at position 501 on spike protein affinity to human ACE2 receptors and if there's a an amino mutation that would cause less/ no affinity to ACE2.