

SARS-CoV-2 ORF8 gene CAA=TAA and AAA=TAA Termination Codon Mutations found mostly in B.1.1.7 Variants was Independent of Popular L84S Point Mutations

Asti Kumar Chakraborty

Department of Biotechnology and Biochemistry, Oriental Institute of Science and Technology, Vidyasagar University, Midnapore 721102, India.

*Corresponding Author

Asti Kumar Chakraborty, Department of Biotechnology and Biochemistry, Oriental Institute of Science and Technology, Vidyasagar University, Midnapore 721102, India, chakraakc@gmail.com

Submitted: 12 Nov 2022; Accepted: 27 Nov 2022; Published: 28 Dec 2022

Citation: Chakraborty, A. K. (2022). SARS-CoV-2 ORF8 gene CAA=TAA and AAA=TAA Termination Codon Mutations found mostly in B.1.1.7 Variants was Independent of Popular L84S Point Mutations. *Int J Clin Med Edu Res.* 1(6), 192-208.

Abstract

Five VOCs of SARS-CoV-2 mainly caused million deaths worldwide and named as B.1.1.7 (U.K.), B.1.351 (South Africa), P.1 (Brazil), B.1.617.2 (India), and B.1.1.529 (Africa). In HIV mediated pathogenesis, small trans activator proteins (TAT, NEF, REV) modulate transcription of cellular genes. Similarly, preliminary reports indicated that corona virus ORF8 protein acts as histone mimics disrupting chromatic structure with many epigenetic changes and immune modulator functions. ORF8 protein had also some similarities to immunoglobulin domains and inhibited HMC-1 and IFN-beta functions.

During evolution a 382-nucleotide deletion ($\Delta 382$) in the ORF8 region of the corona virus genome leads to weak virus load and weak pathogenicity (accession no. MT374101). We BLAST searched deletion boundary and was selected few ORF8 protein truncated mutants. The C>T base change at 27972nt and another A>T base change at 28095nt created two termination codons (CAA=TAA and AAA=TAA) to produce 26AA and 67AA long ORF8 truncated proteins. Similar Blast-N search with oligonucleotides selected at the mutation boundaries gave many ORF8 mutants with distinct S24L, V32L, P38S, R52I, A65V, Y73C, L84S, K92E and V100L mutations with or without TAA termination mutations.

Major mutations found in B.1.1.7 lineage which had spike 69HV and 145Y mutations and ORF1ab polyprotein 3675KSF deletion. However, one ORF8 mutant (accession no. OW221449) belongs to Omicron BA.2 variant with 24LPP spike deletion and others to Omicron BA.5 variants (accession nos. OP733645 and OP671680) with 24LPP and 69HV deletion in the spike protein. One termination codon mutant (accession no. OP711842) has also 63nt ORF7a/b deletions. Mutation did not change the hairpin structure in the ORF8 gene and ORF8 protein formed dimeric stable globular 3-D structure to interact with many host proteins. Clearly, generation of such abundant B.1.1.7 lineage ORF8 protein truncated mutants may be one of the causes for the extinction of Alpha variant of corona virus in 2021. Roles of ORF8 mutants as host proteins modulator were explained in light of other deletions and mutations in corona virus genome.

Keywords: ORF8 protein, SRAS-CoV-2 Mutation, Termination Codon, Truncated ORF8, Host Protein Modulator.

Introduction

The SARS-CoV-2 is the causative agent of the coronavirus disease 2019 (COVID-19), a large global outbreak with severe public health consequences and about million deaths [1]. The MERS-Cove, responsible for several outbreaks since 2012 also has been well characterized. The novel SARS-CoV-2 shares nearly 96% similarity to the bat coronavirus isolate RaTG13, suggesting these

animals are the likely natural reservoir of the virus [2].

Thus, different animal, birds and whale corona viruses were known since 2003 but human corona viruses were appeared in December, 2019 at Wuhan province of China. Within 2 years few thousand variants were sequenced and divided into Alpha, Beta, Delta, Gamma as well as Omicron variants [3]. Analysis suggested

that spike protein (1273AA) of COVID-19 had gone extensive mutations and deletions than large polyprotein ORF1ab (7096aa). The LPP, HV, VYY, FR, L and Y were major deletions occurred in the spike whereas EPE insertion was also reported in Omicron BA.1 variant.

Among the ORF1ab deletions, KSF deletion in nsp1 domain was found only in omicron BA.4 sub-variants and LSG deletion found in omicron BA.1 sub-variant whereas SGF deletion in nsp6 domain was found in most Omicron and Alpha variants [4]. Dominant point mutations D614G and N501Y were important that increased transmission and pathogenicity whereas ~30 mutations mostly in the RBP domain were reported that were not found in deadly B.1.1.7, B.1.617.2 and AY.103 variants. The E484A, T478K, L452R and K417N/T, were very immune-modular. However, Omicron variants were less pathogenic and usually did not require oxygen support and hospitalization unless co-morbidity. Still pneumonia, cough and cold, chest pain, confusion and headache were different symptoms that affected >630 million peoples worldwide.

SARS-CoV-2 is a large positive-stranded RNA virus with ~30000 nucleotides genome. It has structural proteins Membrane (M), Envelope (E), Nucleocapsid (N), Spike (S) coded from 3'-1/3 part of the virus independently but RNA-dependent RNA polymerase was coded from nsp12 domain of ORF1ab polyprotein coded from 2/3 of the 5'-parts of the genome and such polyprotein was degraded into sixteen polypeptides (nsp1-nsp16) [5]. The nsp2 protein is RNA topoisomerase whereas Nsp3 and nsp5 are proteases. The nsp6, nsp7, nsp8, nsp9 and nsp10 were small accessory proteins involved in RNA polymerase replication complex [5].

The nsp14 and nsp15 are nucleases and nsp16 is methyltransferase (C and as well as nsp13 is RNA helicase [6, 7]. Nsp11 is a small peptide and function was not known. ORF3a, ORF6, ORF7a/7b, ORF8 and ORF10 small proteins also coded from 3' end of the genome and have roles in regulating cellular genes [8]. Many drugs were discovered against proteases and RNA polymerases but vaccines were only important remedy that halted the corona virus spread.

Attenuated virus vaccine was developed but engineered spike protein DNA vaccine was very successful whereas mRNA vaccine was also reported. Here, we reported that small ORF8 protein synthesis was hampered due to creation of TAA termination codon in mRNA creating 26AA truncated ORF8 protein. This phenomenon was shown to reduce corona virus load and less severe pathogenicity and likely virus was cleared by host immune system.

Methods

We searched PubMed to get idea on published papers on ORF8 and also searched SARS-CoV-2 NCBI database using BLAST-N and BLAST-X search methods. Multi-alignment of protein was done by

Multiline software [9] and multi-alignment of DNA by CLUSTAL-Omega software (Sievers, F et al., 2011). 1st impression of ORF8 mutants was gained by Blast search of deletion boundary of 120nt sequence and analyzing the sequences with 90-100% similarities.

Blast search of ORF8 full length gene used to get mutant ORF8 proteins with or without termination codon [10]. Then, the other ORF8 mutants were detected by Blast-N search of TAA mutant oligos as well as other oligos selected from point mutation boundaries. Hairpin structure of ORF8 gene 222nt 5'-terminal sequence was done by Oligoaniline 3.1 software (Integrated DNA Technologies). The protein 3-D structure was determined by SWISS-Model software (<https://swissmodel.expasy.org/>) [11].

Results

The genes and proteins in the corona virus (~30kb) and the sequence of ORF8 protein were shown in figure-1. BLAST-N search first selected three ORF8 gene TAA termination mutants (accession nos. MZ212478, MZ294172 and OK234981) using $\Delta 382$ deletion oligo. There was a G>T mutation at 27972nt causing CAA change to 1st TAA termination codon and only 26AA ORF8 protein was produced instead 121 AA. There was a 2nd termination codon created at 68K amino acid codon (AAA=TAA) causing 68AA ORF8 protein.

In accession nos. OP735325, OP735327 and OP735334, there was a C>T mutation at 27964 causing ORF8 S24L mutation. However, there was a creation of another G>T mutation at 28048nt in termination mutants (accession nos. MZ212478, MZ294172 and OK234981) causing R52I AA change. Another Y73C mutation in the ORF8 protein was located nearby 2nd termination codon (figure-2).

Then, we used 1st termination oligo and 2nd termination oligo (Table-1) to select more termination mutants and multi-aligned together (Figure-2A; accession nos. OP711844 and OP711836). The 30kb corona virus genomes alignment was huge (>100 pages) so only the desired portions of the alignment were shown in figures. Figure-2B showed unusual GHVMV deletion (5'-GGTCATGTTATGGTT-3') in the ORF1ab polyprotein in one termination mutant (accession no. OP711844) while all five termination mutants had SGF deletion.

In figure-2C, we demonstrated that all five TAA mutants had HV and Y deletions in the spike protein indicating Alpha (B.1.1.7) variant type. Surely such variant also contained N501Y and D614G dominant mutations in the spike protein required for enhanced transmission and disease severity (data not shown). We further aligned the spike proteins to confirmed the above finding through genetic alignment that indeed 69HV and 145Y deletions occurred (figure-2D).

Table 1: Oligos selected at the TAA termination, point mutation and deletion boundaries used for BLAST-N search

1 st Termination oligo	5'-agtcattgtacttaacatcaacc-3'
2 nd Termination oligo	5'-atgaggctggttcttaacacccattcagt-3'
Δ382 deletion oligo	5'-tctattgtg ctttttagcc ttctgctat tcctgtttt aattatgctt attatctttt gggtctgac gttcgtgtg ttttagatt catctaaacg aacaaactaa aatgtctgat-3'
L84S boundary oligo	5'-tacagtttctgttcacctttacaatta-3'
S24L plus 1st TAA codon oligo	5'-gtttacagttatgtacttaacatcaacca-3'
Y73C plus 2nd TAA codon oligo	5'-atgaggctggttcttaacacccattcagtcacgatatcggt-3'

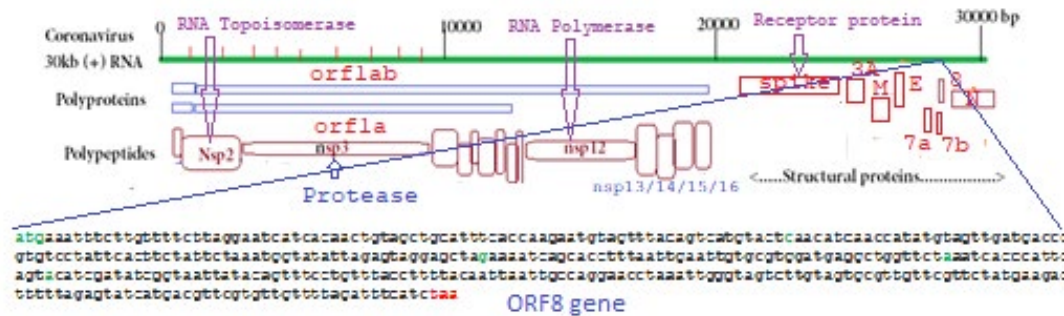


Figure 1: Structure of 29980nt SARS-CoV-2 genome and its protein components. Complete sequence of ORF8 gene was given.

Acc. no. Date of isolation	ORF8 region
MZ213478-12-4-2021	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27999
MZ294172-14-5-2021	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-28000
OK234981-3-6-2021	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27930
OP711836-2nd	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27948
OP711844-1st	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27933
MZ611974-12-3-2021	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27974
OA982176-10-9-2020	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-28020
MW593419-20-5-2020	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27978
B-NC_045512.2-12-2019	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-28020
OP735325-30-12-2020	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27982
OP735325-8-1-2021	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27982
OP735327-30-12-2020	agtcattgtacttaacatcaaccatatgtagttgatgacccgtgtcctattcacttcttatt-27983
//	*** **
MZ213478-12-4-2021	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28119
MZ294172-14-5-2021	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28120
OK234981-3-6-2021	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28050
OP711836-2nd	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28068
OP711844-1st	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28053
MZ611974-12-3-2021	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28094
OA982176-10-9-2020	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28140
MW593419-20-5-2020	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28098
B-NC_045512.2-12-2019	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28140
OP735325-30-12-2020	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28102
OP735325-8-1-2021	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28102
OP735327-30-12-2020	atgaggctggttcttaacacccattcagtcacgatatcggttaattatacagtttccct-28103
	*** **

Figure 2A: Multi-alignment of COVID-19 sequences demonstrating ORF8 gene TAA 1st termination codon (CAA=TAA) of which four have also 2nd termination codon (AAA=TAA).

Acc. no. Date of isolation	ORF1ab region
MZ213478-12-4-2021	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-537
MZ294172-14-5-2021	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-538
OK234981-3-6-2021	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-468
OP711836-2nd	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-486
OP711844-1st	acgttcggatgctcgaactgcacctcat---GHVMV-----gagctggtagcagaact-471
MZ611974-12-3-2021	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-501
OA982176-10-9-2020	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-540
MW593419-20-5-2020	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-498
B-NC_045512.2-12-2019	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-540
OP735325-30-12-2020	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-502
OP735325-8-1-2021	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-502
OP735327-30-12-2020	acgttcggatgctcgaactgcacctcatgggtcatgttatgggttagctggtagcagaact-503

MZ213478-12-4-2021	tagtttg---SGF-----aagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11328
MZ294172-14-5-2021	tagtttg-----aagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11329
OK234981-3-6-2021	tagtttg-----aagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11259
OP711836-2nd	tagtttg-----aagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11277
OP711844-1st	tagtttg-----aagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11262
MZ611974-12-3-2021	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11301
OA982176-10-9-2020	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11340
MW593419-20-5-2020	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11298
B-NC_045512.2-12-2019	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11340
OP735325-30-12-2020	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11302
OP735325-8-1-2021	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11302
OP735327-30-12-2020	tagtttgctcgtggttttaagctaaaagactgtgttatgtatgcatcagctgttagtggttact-11303

Figure 2B: Deletion in the ORF1ab protein in few ORF8 TAA termination codon mutants.

Acc. No. Date of isolation	Spike protein region
MZ213478-12-4-2021	cttggttccttacctttcttttccaatgttacttgggtccatgctatc---HV---tctggggac-21762
MZ294172-14-5-2021	cttggttccttacctttcttttccaatgttacttgggtccatgctatc-----tctggggac-21763
OK234981-3-6-2021	cttggttccttacctttcttttccaatgttacttgggtccatgctatc-----tctggggac-21693
OP711836-2nd TAA	cttggttccttacctttcttttccaatgttacttgggtccatgctatc-----tctggggac-21711
OP711844-1st TAA	cttggttccttacctttcttttccaatgttacttgggtccatgctatc-----tctggggac-21696
MZ611974-12-3-2021	cttggttccttacctttcttttccaatgttacttgggtccatgctatc-----tctggggac-21735
OA982176-10-9-2020	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21780
MW593419-20-5-2020	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21738
B-NC_045512.2-12-2019	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21780
OP735325-30-12-2020	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21742
OP735325-8-1-2021	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21742
OP735327-30-12-2020	cttggttccttacctttcttttccaatgttacttgggtccatgctatcacatgtctctggggac-21743

MZ213478-12-4-2021	tcaattttgtaatgatccatttttgggtgtt---Y---taccacaaaaaacacaaaagtggat-21999
MZ294172-14-5-2021	tcaattttgtaatgatccatttttgggtgtt---taccacaaaaaacacaaaagtggat-22000
OK234981-3-6-2021	tcaattttgtaatgatccatttttgggtgtt---taccacaaaaaacacaaaagtggat-21930
OP711836-2nd	tcaattttgtaatgatccatttttgggtgtt---taccacaaaaaacacaaaagtggat-21948
OP711844-1st	tcaattttgtaatgatccatttttgggtgtt---taccacaaaaaacacaaaagtggat-21933
MZ611974-12-3-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-21975
OA982176-10-9-2020	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-22020
MW593419-20-5-2020	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-21978
B-NC_045512.2-12-2019	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-22020
OP735325-30-12-2020	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-21982
OP735325-8-1-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-21982
OP735327-30-12-2020	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaaacacaaaagtggat-21983

Figure 2C: Multi-alignment of SARS-CoV-2 sequences as in (figure-2A) to show the deletion in the spike protein of ORF8 termination codon mutants. Deletion of 69HV and 145Y indicated that ORF8 mutants were B.1.1.7 variant.

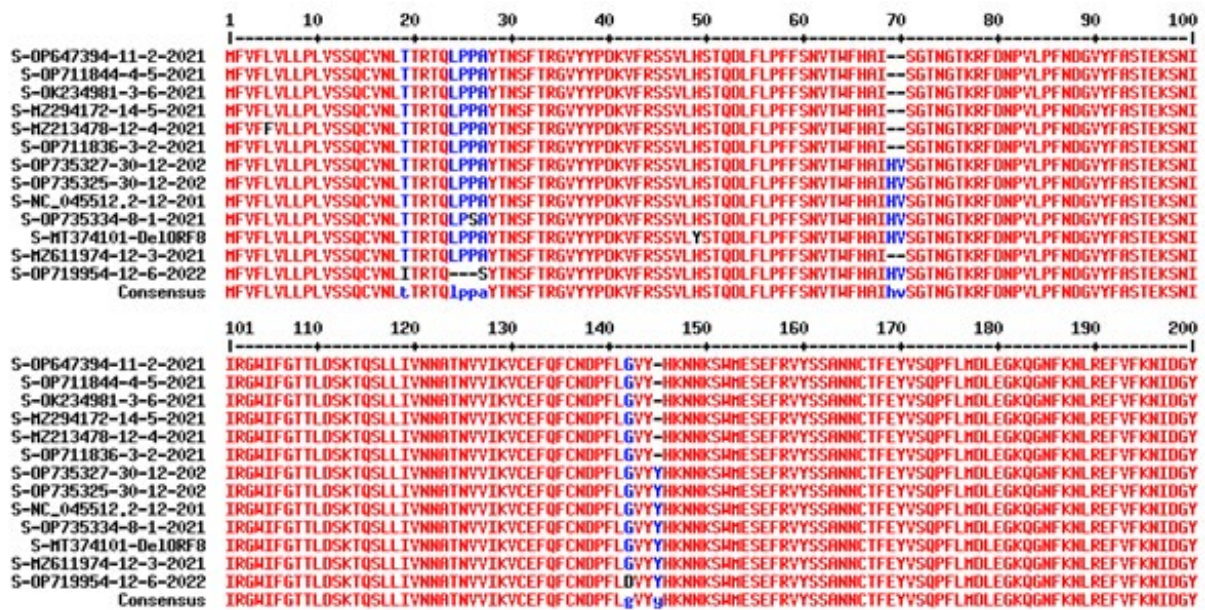


Figure 2D: Multi-alignment of Spike proteins of ORF8 mutants as discussed above. The LPP deletion but not HV deletion indicated accession no. OP719954 belong to an Omicron BA.2 variant. The 69HV and 145Y deletions indicated B.1.1.7 lineage.

Figure-3 demonstrated the presence of R52I (AGA=ATA) mutation in ORF8 termination codon mutants while in figure-4 similar Y73C mutation (TAC=TGC) located after 2nd TAA termination codon was presented (accession nos. MZ213478, MZ294172 and OK234981). In figure-5, we also showed the S24L amino acid change in ORF8 gene without creating TAA termination codons. But in other search we found S24L mutants with creating ORF8 1st termination codon (see, figure-7B). Such findings were done by BLAST-X search of 60nt sequence with distinct mutation selected through multi-alignment.

Next, we demonstrated the abundant ORF8 protein L84S mutants as well as one 119DF two amino acids deletion mutant without generating TAA termination codons (figure-6A). No 1st (not shown here) and 2nd (boxed) TAA termination codons were found in ORF8 gene L84S mutants. These sequences were selected through BLAST-N search using L84S oligo (Table-1). Further, we confirmed the variant type of such mutants demonstrating ORF1ab and spike proteins deletions (figure-6B). The data indicated the presence of LPP and HV deletions in the spike as well as 31ERS deletion in the nucleocapsid (N) protein of corona virus confirming Omicron BA.5 variants (accession nos. OP733645, OP671680).

While accession no. OW221449 had only 24LPP deletion in the spike plus 31ERS deletion in the N protein being an Omicron BA.2 variant. Note that no 141KSF ORF1ab protein deletion was found which was a character of Omicron BA.4 variant. Neither EPE insertion in the spike was found which was a confirmation of Omicron BA.1 variant. Surprisingly, three Omicron variants ORF8 L84S mutants had 26nt 3'UTR deletion and characterization of such deletion never been discussed before [12]. The other accession numbers had no 157FR deletion in the spike and such variant were earlier variants like Beta and P.1 with D614G mutation.

However, Δ382 deletion mutant used to select first TAA mutants had D614 indicating primitive (2020) variant and such big deletion in the ORF8 gene was limited!

Next, we selected more S24L ORF8 mutations with 1st termination codon oligo plus S24L oligo by BLAST-N search (figure-7). In figure-7A we showed the alignment of many accession numbers with 100% homology and 100% cover with the oligo and many 1st TAA mutants of the ORF8 gene selected (figure-7B) whereas no mutation in the 2nd termination area was found. Given to the fact, we yet to get a 2nd termination mutant only without 1st TAA mutation.

Because, presence of 1st TAA created only 26AA ORF8 protein and need for the 2nd TAA not required then. It may happen that there is a sub-genomic initiation site upfront of 2nd termination site creating more smaller spliced ORF8 mRNA and proteins [13].

In figure-7C we confirmed that among the many S24L ORF8 mutants accession numbers ON113700 and OU200773 were B.1.1.7 lineage with spike HV and 1Y deletions as well as SGF ORF1ab protein deletion. But no Delta variant and Omicron variant were detected in this search and thus other accession numbers might be primitive corona virus isolates. As we continued our search for TAA termination ORF8 mutants, we used 2nd TAA termination codon Y73C oligo. It appeared again that although we used 2nd termination oligo, we selected both 1st and 2nd TAA termination codons (figure-8). All B.1.1.7 variants were selected with SGF deletion in ORF1ab protein as well as 69HV+145Y spike deletions (data not shown).

Only OP711842 variant has 63nt deletion in ORF7a/7b region. Our attempts to get mutant with only ORF8 2nd termination codon was failed. We do not understand what is the necessary of 2nd termination codon as active 1st termination codon means no ORF8

protein generates. We found that search with R52I mutation oligo could not generated with a termination codon mutant. Further, we used full length ORF8 gene (normal vs TAA mutated plus point mutation vs TAA repaired plus point mutation) to BLAST-X search and identified few mutant ORF8 alleles which demonstrated in Figure-9A and figure-9B. In this search, we got V32L, P38S, R52I, A65V, Y73C, K92E and V100L mutations in the ORF8 gene.

Then, we aligned the ORF8 mutant protein as shown in Figure-10A and a phylogenetic analysis shown in figure-10B. Similarly, we addressed the RNA structure in the ORF8 gene region and found that no basic changes in hairpin structure was occurred between normal and TAA mutated RNAs (figure-11). We further addressed the overall 3-D protein structures of mutated ORF8 protein but with limitation, a compact folding was found as compared to

animal corona virus ORF8 protein available (PDB ID:7F8I). Ramachandran plot also confirmed the alpha-helical globular structure of the ORF8 protein. Such compact folded structure of ORF8 protein was necessary to interact with host proteins.

In summary, our search generated many ORF8 TAA 1st termination codon plus TAA 2nd termination codon mutants (accession numbers, MZ213478, MZ394172, OK234981, OP711836, OP711844, OP683545, OP711842, OP711837, OX020484, OP585720 and OW999874). Similarly, we demonstrated few only 1st TAA termination codon mutants (accession numbers, MZ497707, ON113700, OU200773, MW98643, OP735332 and ON084987). This confirmed that such ORF8 mutation created no ORF8 protein and likely reduced the severity of corona virus mediated pathogenesis.



Figure 3: Detection of R52I (AGA=ATA) mutation in ORF8 termination codon mutants (accession nos. MZ213478, MZ294172 and OK234981).

Acc.no.Date of isolation ORF8 2nd TAA region Y73C

MZ213478-12-4-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28119

MZ294172-14-5-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28120

OK234981-3-6-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28050

MZ611974-12-3-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28094

OA982176-10-9-2020 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28140

(A) MW593419-20-5-2020 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28098

B-NC_045512.2-12-2019 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28140

OP735325-30-12-2020 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28102

OP735324-8-1-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28102

OP735327-30-12-2020 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct-28103

MZ213478-12-4-2021 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct 28119

blastX-ORF8

ORF8 protein [Severe acute respiratory syndrome coronavirus 2] Sequence ID: [QVU06321.1](#) Length: 121.

Score	Expect	Method	Identities	Positives
39.3 bits(90)	0.030	Composition-based stats.	18/19(95%)	18/19(94%)

(B) Query 3 EAGS*SPIQCIDIGNYTVS 59

EAGS SPIQCIDIGNYTVS

Sbjct 64 EAGSKSPIQCIDIGNYTVS 82

B-NC_045512.2-12-2019 atgaggctgggtcttaatacaccattcagtgcatcgatatacggttaattatacagtttcct 28140

blastX-Wuhan

ORF8 protein, partial [Severe acute respiratory syndrome coronavirus 2] Sequence ID: [UHR93329.1](#) Length: 90.

Score	Expect	Method	Identities	Positives
43.9 bits(102)	3e-04	Composition-based stats.	19/19(100%)	19/19(100%)

Query 3 EAGSKSPIQYIDIGNYTVS 59

EAGSKSPIQYIDIGNYTVS

Sbjct 64 EAGSKSPIQYIDIGNYTVS 82

Figure 4: Demonstration of ORF8 gene Y73C mutation (TAC=TCG) located after 2nd TAA termination codon. (A) Multi-alignment of ORF8 mutants and (B) BlastX search to get mutant AA.

Acc.no.Date of isolation 2nd termination codon ORF8 protein region

MZ213478-12-4-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27999

MZ294172-14-5-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt28000

OK234981-3-6-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27990

MZ611974-12-3-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27974

OA982176-10-9-2020 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt28020

(A) MW593419-20-5-2020 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27978

B-NC_045512.2-12-2019 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt28020

OP735325-30-12-2020 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27982

OP735324-8-1-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27982

OP735327-30-12-2020 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27982

MZ213478-12-4-2021 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27999

blastX-ORF8

ORF8 protein, partial [Severe acute respiratory syndrome coronavirus 2] Sequence ID: [UAM18414.1](#) Length: 55.

Score	Expect	Method	Identities	Positives
(B) 42.0 bits(97)	0.001	Compositional matrix adjust.	18/19(95%)	18/19(94%)

Query 3 SCT*HQPYVWDDPCPIHFY 59

SCT HQPYVWDDPCPIHFY

Sbjct 24 SCTQHQPYYVDDPCPIHFY 42

OP735327-30-12-2020 agtcatgtacttaacatcaaccatatagtagttgatgacccgtgtcctattcacttctatt27982

blastX-ORF8

ORF8 protein, partial [Severe acute respiratory syndrome coronavirus 2] Sequence ID: [UDH89230.1](#) Length: 84.

Score	Expect	Method	Identities	Positives
(C) 45.4 bits(106)	8e-05	Compositional matrix adjust.	19/19(100%)	19/19(100%)

Query 3 LCTQHQPYYVWDDPCPIHFY 59

LCTQHQPYYVWDDPCPIHFY

Sbjct 24 LCTQHQPYYVWDDPCPIHFY 42

Figure 5: Demonstration of S24L amino acid change in ORF8 mutant without creating TAA termination codons. But in other search we found S24L mutants with creating ORF8 1st termination codon (see, figure-7B).

Acc no. Date of virus isolation			
OW221449-8-3-2022	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28122	<div style="border: 1px solid black; padding: 2px; display: inline-block;"> ORF8 protein region </div>
OP732645-4-10-2022	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28114	
OP671680-5-10-2022	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28077	
OP581726-3-2-2021	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28077	
OP471800-2-6-3-2020	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28086	
ON532661-15-2-2022	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28092	
ON653598-12-9-2020	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28129	
OW663380-6-2-2022	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28129	
NC_045512.2-12-2019	atgaggctggttctaaatcaccattcagttacatcgatatcggttaattatacagtttcct	28140	
	*****2nd termination zone*****		
OW221449-8-3-2022	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28182	<div style="border: 1px solid black; padding: 2px; display: inline-block;"> 2nd termination zone </div>
OP732645-4-10-2022	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28174	
OP671680-5-10-2022	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28137	
OP581726-3-2-2021	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28137	
OP471800-2-6-3-2020	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28146	
ON532661-15-2-2022	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28152	
ON653598-12-9-2020	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28189	
OW663380-6-2-2022	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28189	
NC_045512.2-12-2019	gttcacctttttacaatttaattgcccagggaacctaaattgggtagtccttctagtcgctgtgtt	28200	
	*****L84S*****		
OW221449-8-3-2022	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28242	<div style="border: 1px solid black; padding: 2px; display: inline-block;"> L84S </div>
OP732645-4-10-2022	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28234	
OP671680-5-10-2022	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28197	
OP581726-3-2-2021	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28191	
OP471800-2-6-3-2020	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28206	
ON532661-15-2-2022	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28212	
ON653598-12-9-2020	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28249	
OW663380-6-2-2022	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28249	
NC_045512.2-12-2019	cggttctatgaagacttttttagagttatcatgacgttcgtgtgttttagatttcacataaaa	28260	
	*****ORF8 termination codon*****		

Figure 6A: Demonstration of abundant ORF8 protein L84S mutants as well as one 119DF deletion mutant without generating TAA termination codons. No 1st termination codon (not shown here) and 2nd termination codon (boxed) were found in ORF8 gene L84S mutants.

OW221449-8-3-2022	tagttttg-----aaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11331
OP732645-4-10-2022	tagttttg---8GF---aaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11329
OP671680-5-10-2022	tagttttg-----aaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11329
OP581726-3-2-2021	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11286
OP471800-2-6-3-2020	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11286
ON532661-15-2-2022	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11292
ON653598-12-9-2020	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11329
OW663380-6-2-2022	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11329
NC_045512.2-12-2019	tagttttgctgggttttaagctaaagactgtgttatgtatgcatcagctgtagtggttact	11340
	*****L84S*****	
OW221449-8-3-2022	tcagtggtgtaattcttataaccagaactcaat-----catabactaattcttttcac	21642
OP732645-4-10-2022	tcagtggtgtaattcttataaccagaactcaat---LFP---catabactaattcttttcac	21640
OP671680-5-10-2022	tcagtggtgtaattcttataaccagaactcaat-----catabactaattcttttcac	21603
OP581726-3-2-2021	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21606
OP471800-2-6-3-2020	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21606
ON532661-15-2-2022	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21612
ON653598-12-9-2020	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21649
OW663380-6-2-2022	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21649
NC_045512.2-12-2019	tcagtggtgtaattcttataaccagaactcaattacccctgcatacactaattcttttcac	21660
	*****L84S*****	
OW221449-8-3-2022	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21762
OP732645-4-10-2022	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21754
OP671680-5-10-2022	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21717
OP581726-3-2-2021	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21726
OP471800-2-6-3-2020	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21726
ON532661-15-2-2022	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21732
ON653598-12-9-2020	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21769
OW663380-6-2-2022	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21769
NC_045512.2-12-2019	cttggttcttaacctttctttccaaatgttacttgggttccatgctatcacatgtctctgggac	21780
	*****L84S*****	
OW221449-8-3-2022	gtttggtggaccctcagattcaactggcagtaaacagaatg---ERS---gtggggcgcg	28353
OP732645-4-10-2022	gtttggtggaccctcagattcaactggcagtaaacagaatg---gtggggcgcg	28345
OP671680-5-10-2022	gtttggtggaccctcagattcaactggcagtaaacagaatg---gtggggcgcg	28308
OP581726-3-2-2021	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28311
OP471800-2-6-3-2020	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28326
ON532661-15-2-2022	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28332
ON653598-12-9-2020	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28369
OW663380-6-2-2022	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28369
NC_045512.2-12-2019	gtttggtggaccctcagattcaactggcagtaaacagaatggagaaacgcagtgggggcgcg	28380
	*****L84S*****	
OW221449-8-3-2022	gggaggagcttgaaagagccaccatttttcacc-----3' UTR-----t	29707
OP732645-4-10-2022	gggaggagcttgaaagagccaccatttttcacc-----t	29699
OP671680-5-10-2022	gggaggagcttgaaagagccaccatttttcacc-----t	29662
OP581726-3-2-2021	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29691
OP471800-2-6-3-2020	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29706
ON532661-15-2-2022	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29712
ON653598-12-9-2020	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29749
OW663380-6-2-2022	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29749
NC_045512.2-12-2019	gggaggagcttgaaagagccaccatttttcaccggggcaccagggagtagcagtcagagtg	29760

Figure 6B: Demonstration of differential ORF1ab, S, N and 3'-UTR deletions in ORF8 L84S mutants. The accession nos. OW221449, OP732645, OP671680 were Omicron variant with ERS N-protein deletion. The OW221449 is omicron BA.2 with no 69HV deletion whereas both OP732645 and OP671680 are omicron BA.5 as no 141KSF deletion for BA.4 variant and no EPE insertion for BA.1 variant. However, we did not find B.1.1.7 variant here.

Sequences producing significant alignments									
Download Select columns Show 100									
select all 100 sequences selected									
GenBank Graphics Distance tree of results MSA Viewer									
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/1650757/2021 ORF1ab poly...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19778	ON113700.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/19349/2020 ORF1ab polyse...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19807	ON084987.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/hCoV_19_USA_ID_IVREF_7...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19819	OX566688.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/MN-CDC-BX264153753205...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19864	MZ717946.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/WY119558/2021 ORF1ab e...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19889	MZ497707.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/NV NSPHL-1V_308458/202...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19782	MZ473687.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_chromosome_1	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19854	OU200773.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_chromosome_1	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19854	OU201727.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/IL-CDC-1C0070592/2021 O...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19546	MZ412275.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_chromosome_1	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19884	OX055692.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_complete genome: monoonartite	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19884	OX0556978.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CO-CDC-MMR016889131/20...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19782	MZ298065.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_chromosome_1	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19884	OU0334614.1
✓	Severe acute respiratory syndrome coronavirus 2 genome assembly_chromosome_1	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19884	OU0334612.1
✓	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/TX-CDC-ASC210037495/20...	Severe acute res...	58.0	58.0	100%	2e-05	100.00%	19858	MZ282053.1

Figure 7A: Selection of more S24L ORF8 mutation with 1st termination codon oligo by BLAST-N search.

Acc. no. Date of isolation	S24L 1st termination codon generated in most	
MZ497707-24-4-2021	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	28008
ON113700-30-3-2021	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	27963
OU200773-3-5-2021	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	27972
MW731336-11-2-2021	agt y atgtact y aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	28017
MW986843=8-3-2021	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	27990
OP735332-8-1-2021	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	27982
ON084987-7-10-2020	agt t atgtact t aacatcaaccatagtgtgatgacccgtgtcctattcacttctatt	27984
NC_045512.2-12-2019	agt c atgtact c aaatcaaccatagtgtgatgacccgtgtcctattcacttctatt	28020

MZ497707-24-4-2021	ctaaatgggtatattagagtaggagctag a aaatcagcacctttaaattgaattgtgcgtgg	28068
ON113700-30-3-2021	ctaaatgggtatattagagtaggagct a taaaatcagcacctttaaattgaattgtgcgtgg	28023
OU200773-3-5-2021	ctaaatgggtatattagagtaggagct a taaaatcagcacctttaaattgaattgtgcgtgg	28032
MW731336-11-2-2021	ctaaatgggtatattagagtaggagctak a aaatcagcacctttaaattgaattgtgcgtgg	28077
MW986843=8-3-2021	ctaaatgggtatattagagtaggagctag a aaatcagcacctttaaattgaattgtgcgtgg	28050
OP735332-8-1-2021	ctaaatgggtatattagagtaggagctag a aaatcagcacctttaaattgaattgtgcgtgg	28042
ON084987-7-10-2020	ctaaatgggtatattagagtaggagctag a aaatcagcacctttaaattgaattgtgcgtgg	28044
NC_045512.2-12-2019	ctaaatgggtatattagagtaggagctag a aaatcagcacctttaaattgaattgtgcgtgg	28080

Region of 2nd termination codon but no A>T mutation here		
MZ497707-24-4-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28128
ON113700-30-3-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28083
OU200773-3-5-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28092
MW731336-11-2-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28137
MW986843=8-3-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28110
OP735332-8-1-2021	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28102
ON084987-7-10-2020	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28104
NC_045512.2-12-2019	atgaggctggttct t aaatcaccattcagtagcatcgatcggtaattatacagtttctt	28140

MZ497707-24-4-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28248
ON113700-30-3-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28203
OU200773-3-5-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28212
MW731336-11-2-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28257
MW986843=8-3-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28230
OP735332-8-1-2021	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28222
ON084987-7-10-2020	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28224
NC_045512.2-12-2019	cgttctatgaagactttttagagtatcatgacgttcgtgttgttttagattttatct aaa	28260

ORF8 termination codon

Figure 7B: Selection of more S24L ORF8 mutant with 1st TAA termination codon oligo but no 2nd termination codon was generated here.

Acc.no. Date of isolation	ORF1ab protein region
MZ497707-24-4-2021	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11328
ON113700-30-3-2021	tagtttg---SGF---aagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11292
OU200773-3-5-2021	tagtttg-----aagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11301
MW731336-11-2-2021	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11337
MW986843-8-3-2021	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11310
OP735332-8-1-2021	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11302
ON084987-7-10-2020	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11304
NC_045512.2-12-2019	tagtttgtctggttttaagctaaaagactgtgttatgtatgcatcagctgtagtgttact 11340
//	*****
	Spike protein region
MZ497707-24-4-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctgggac 21768
ON113700-30-3-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatc-HV---tctgggac 21726
OU200773-3-5-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatc-----tctgggac 21735
MW731336-11-2-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctgggac 21777
MW986843-8-3-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctgggac 21750
OP735332-8-1-2021	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctaggac 21742
ON084987-7-10-2020	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctgggac 21744
NC_045512.2-12-2019	cttggttcttacctttctttttccaatgttacttggttccatgctatacatgtctctgggac 21780

MZ497707-24-4-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 22008
ON113700-30-3-2021	tcaattttgtaatgatccatttttgggtgtt-Y-taccacaaaaacaacaaaagtggat 21963
OU200773-3-5-2021	tcaattttgtaatgatccatttttgggtgtt---taccacaaaaacaacaaaagtggat 21972
MW731336-11-2-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 22017
MW986843-8-3-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 21990
OP735332-8-1-2021	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 21982
ON084987-7-10-2020	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 21984
NC_045512.2-12-2019	tcaattttgtaatgatccatttttgggtgtttattaccacaaaaacaacaaaagtggat 22020

Figure 7C: Deletions in ORF1ab and spike proteins of S24L mutants to demonstrate few ORF8 S24L mutants are Alpha variant (accession nos. ON113700 and OU200773) but not Delta or Omicron variants.

Acc.no. Date of isolation	ORF8 gene region of SARS-CoV-2
NC_045512.2-12-2019	agtcattgtactcaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 28020
OP683545-27-6-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 27963
OP711842-13-4-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 27884
OP711837-3-2-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 27948
OX020484-24-4-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 28002
OP585720-1-1-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 27948
OW999874-26-2-2021	agtcattgtacttaacatcaaccatattgtagttgatgacccgtgtcctattcacttctatt 28002

	Termination codon
NC_045512.2-12-2019	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28080
OP683545-27-6-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28023
OP711842-13-4-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 27944
OP711837-3-2-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28008
OX020484-24-4-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28062
OP585720-1-1-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28008
OW999874-26-2-2021	ctaaatggtatattagagtaggagctatagaaaatcagcacctttaattgaattgtgcgtgg 28062

	Point mutation and no termination codon
NC_045512.2-12-2019	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28140
OP683545-27-6-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28083
OP711842-13-4-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28004
OP711837-3-2-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28068
OX020484-24-4-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28122
OP585720-1-1-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28068
OW999874-26-2-2021	atgaggctggttctaaatcaccattcagtgcatcgatatcggttaattatacagtttct 28122

	Termination codon point mutation

Figure 8: Selection of ORF8 mutants using 2nd TAA termination codon oligo with Y73C mutation. All B.1.1.7 variants were selected with 3675SGF deletion in ORF1ab protein as well as 69HV+145Y spike deletions (data not shown). Only OP711842 variant has 63nt deletion in ORF7a/7b region. We found that R52I mutation could not generated with a termination codon.

(A) D-NC_045512.2-12-2019 (Wuhan corona virus ORF8 gene)
 atgaaattttctgttttcttaggaatcacaactgtaggtgcatttcaccaagaatgtagtttacagtcagtcactcaacatcaaccatagtagttgatgaccc
 gtgtcttattcacttctattctaaatggtatattagagtagggagctagaaaatcagcacctttaattgaattgtgctggatgaggtggttctcaatcaccattc
 agtcacatgatatcggttaattatcacagtttctctgtttaccttttacaattaattgacagggaacctaaattgggttagtcttgttagtgctgttctctatgaagac
 tttttagagatcatgacgttcgtgtgttttagatttcattc**ttaa**

ORF8 protein [SARS-CoV-2, Accession no. NC_045512.2] Sequence ID: [YP_009724396.1](#) Length: 121.

Score	Expect	Method	Identities	Positives
251 bits(642)	3e-84	Compositional matrix adjust.	121/121(100%)	121/121(100%)
Query 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	100	
Sbjct 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	60	
Query 101	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	361	
Sbjct 61	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	121	

(B) OR234981-3-6-2021 (ORF8 gene with two TAA termination codons)
 atgaaattttctgttttcttaggaatcacaactgtaggtgcatttcaccaagaatgtagtttacagtcagtcactcaacatcaaccatagtagttgatgaccc
 gtgtcttattcacttctattctaaatggtatattagagtagggagctagaaaatcagcacctttaattgaattgtgctggatgaggtggttctcaatcaccattc
 agtcacatgatatcggttaattatcacagtttctctgtttaccttttacaattaattgacagggaacctaaattgggttagtcttgttagtgctgttctctatgaagac
 tttttagagatcatgacgttcgtgtgttttagatttcattc**ttaa**

ORF8 protein [SARS-CoV-2, Accession no. OL739852] Sequence ID: [UFX85754.1](#) Length: 121.

Score	Expect	Method	Identities	Positives
243 bits(619)	1e-80	Compositional matrix adjust.	119/121(98%)	119/121(98%)
Query 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	100	
Sbjct 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	60	
Query 101	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	361	
Sbjct 61	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	121	

ORF8 protein [SARS-CoV-2, Accession no. MZ013607] Sequence ID: [QUE21848.1](#) Length: 121.

Score	Expect	Method	Identities	Positives
239 bits(610)	3e-79	Compositional matrix adjust.	117/121(97%)	118/121(97%)
Query 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	100	
Sbjct 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	60	
Query 101	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	361	
Sbjct 61	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	121	

ORF8 protein [SARS-CoV-2, Accession no. MZ341638] Sequence ID: [QWE64732.1](#) Length: 121.

Score	Expect	Method	Identities	Positives
239 bits(610)	3e-79	Compositional matrix adjust.	117/121(97%)	118/121(97%)
Query 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	100	
Sbjct 1	MRFLVFLGIITTVAAAFHQECSLQSC	TQHQPYYVDDPCPIHFYSEWYIRVGARKSAPLIEL	60	
Query 101	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	361	
Sbjct 61	CVDEAGS	SPIQCIDIGNYTVSCLPFTINQEPKLGSLVVRCSFYEDFLEYHMDVRVVLDFI	121	

Figure 9A: Demonstration of new ORF8 mutation by BLAST-X search of full length ORF8 gene. (A) Wuhan corona virus ORF8 gene analysis and (B) two TAA termination ORF8 gene analysis.



Figure 9B: Selection more ORF8 mutants by BLAST-X search using mutant ORF gene but repairing two TAA mutations. Here, MW644355 mutant may be B.1.1.7 lineage.

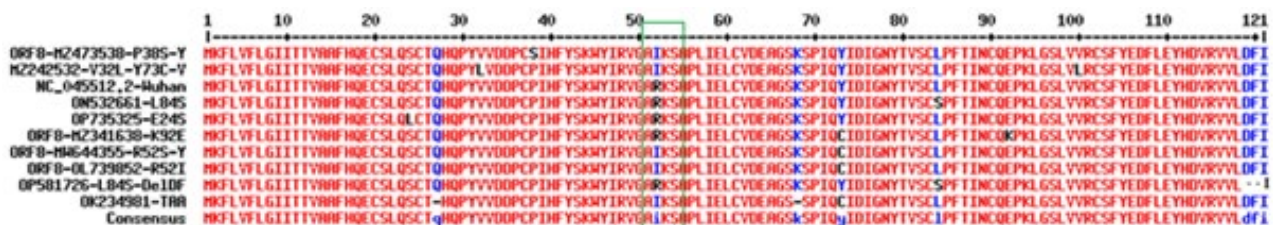


Figure 10A: Multi-alignment of mutated different ORF8 proteins of COVID-19. However, “ARKS” domain (green boxed) similarities with histones was abolished in R52I mutants lowering the interaction with host chromatin (Kee, J et al., 2022). This could be stabilize the corona virus load increasing virus severity.

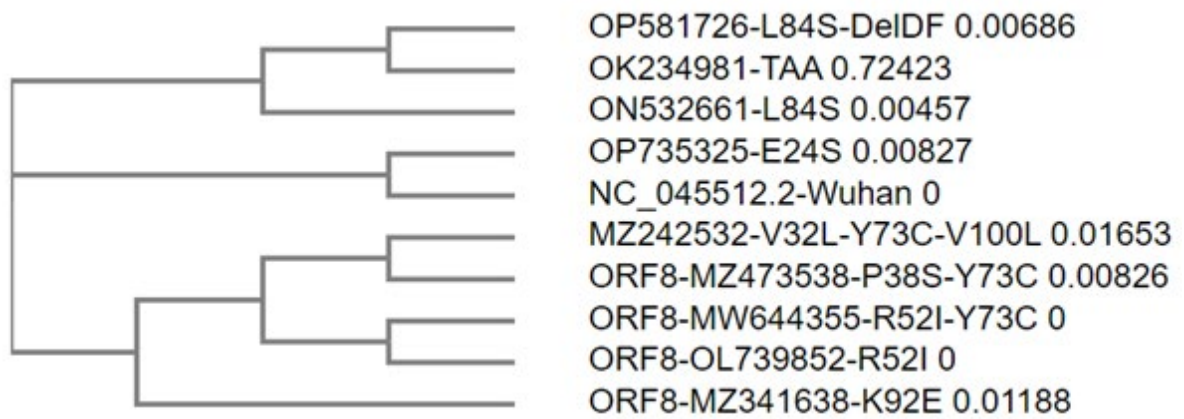


Figure 10B: A phylogenetic tree of corona virus ORF protein mutants.

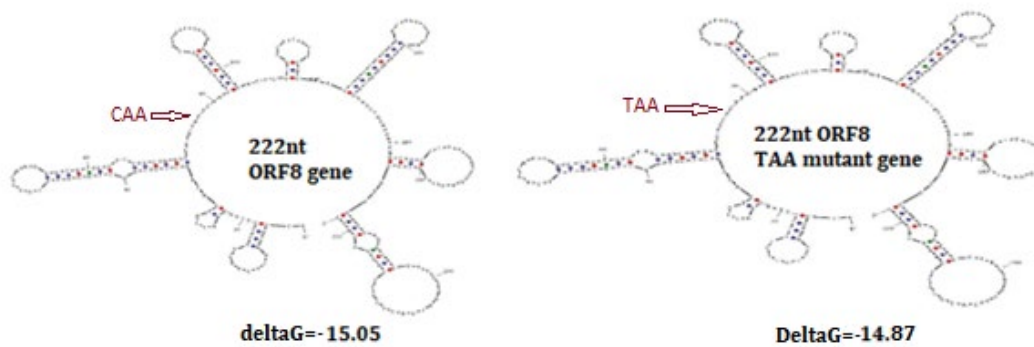


Figure 11: Demonstration of identical hairpin structure of 222nt ORF8 gene of normal and termination mutant of SARS-CoV-2.

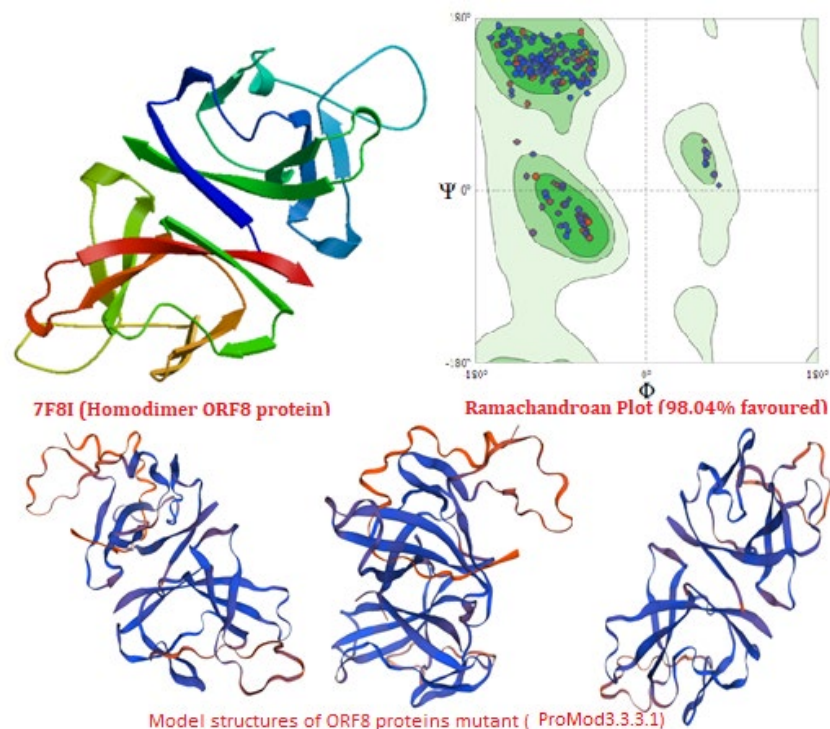


Figure 12: Model structures of Mutated ORF8 proteins (below) and a model of bat ORF8 protein (7F8I) with 98% similarity (upper) and Ramachandran plot to show stable structure of ORF8 dimer (SWISS-Model)

Discussion

Over time, the coronavirus has undergone mutations and deletions and different variants reported in different parts of the world with different time since December, 2019. These variants and mutants are Alpha (B.1.1.7), B.1.177, B.1.258, Beta (B.1.351), Gamma (P.1/B.1.1.28.1), Delta (B.1.617.2), Kappa (B.1.617.1), Epsilon (B.1.427/B.1.429), Zeta (P.2); Eta (B.1.525), B.1.1.20, Iota (B.1.526) and B.1.1.298 (Mink Variant). These mutations are of major health concerns, as they reportedly exacerbate the infectious rate of the virus [4]. We found here major ORF8 truncated mutants were B.1.1.7 lineage. Interestingly, RNA recombination generated few omicron lineages with ORF8 protein TAA termination codon mutants.

Initially we selected few TAA termination mutants (figure-2) but BLAST-N search with mutant oligos selected many of such mutants. Whereas, BLAST-X search with mutated full length ORF8 gene contributed more such mutants to characterize. In truth hundred of TAA termination codon mutants detected in the NCBI Database but many of them were partially sequenced and we avoided such sequences in our multi-alignment data (see, figure-6A; figure-7B). RNA sequencing was done extensively and Delta, BA.2.75, BA.4.6, BA.5.2.1, BF.7 and BE.1.1 were predominantly circulating now. We have attempted to characterize the 121 AA long ORF8 protein in the database and such protein was not expressed in many corona virus variants[14].

The ORF8 L84S mutant was more stable and more soluble than wild type and in higher temperature it existed as aggregated conformation. Moreover, ORF8 precipitated at acidic pH and this precipitation was reversed when the solution pH was shifted to neutral [15]. However, we did not find TAA termination codon mutant in L84S mutants but L24S mutants. Preliminary study indicated few dozen proteins like PVR, IRF3, ATF6, Belin 1, FK506-binding protein 10, EDEM, vitronectin, OPJ94, Sec62163, VIP36, TRFT3 and PLAT etc interacted with ORF8 protein regulating protein folding, apoptosis and interferon production. Such process likely favours COVID-19 survival in host cells inhibiting immune control mechanisms [16]. Genetic analysis pointed a severe deletion in ORF8 (Δ 382) caused less severe corona infections likely due to low viral load and increased immune clearance [17-20]. However, in cell culture study with such deletion mutant contradicted the finding of lower viral load with no change of cellular transcriptional profile [21].

A recent study suggested that ORF8 mediates immune evasion by downregulating MHC-I molecules like HLA-A2 [22,23] and such finding clearly established the IgG domains similarity of ORF8 to modulate host immune functions and chromatin structure [24]. Dominant mutation of small regulatory protein ORF7a limits viral suppression of the interferon response [25] and we detected one such 62nt deletion abolishing the production of both ORF7a and ORF7b proteins (see, accession no. OP711842).

The ORF8 protein also activates the IL-17 and NLRP3 signalling

pathways, upregulates proinflammatory factors and inhibits the production of IFN types I and II. The study also investigated the ability of compounds isolated from *C. volubilis* leaves to interact with the initiation and termination codons of ORF8 mRNAs isolated from the whole genome of SARS-CoV-2 using computational tools.

The ORF8 mRNA sequences of the translation initiation sites and translation termination sites encoding ORF8 amino acids were retrieved from the full genome of SARS-CoV-2. Molecular docking studies revealed strong molecular interactions of the isolated compounds with the ORF8 mRNA. The immunomodulatory potentials of the isolated compounds were investigated on neutrophil phagocytic respiratory bursts using luminol-amplified chemiluminescence technique [26].

There are also reports that ORF8 modulates vesicular traffic through the unfolded protein response and, therefore, ER stress by stimulating the ATF6 and IRE1 pathways through the upregulation of the GRP78 and GRP94 chaperones [27]. Another small corona virus protein ORF9b interacted with Tom70 mitochondrial chaperone as revealed by cryo-electron microscopy [28]. Similarly, study indicated ORF3a also has trans activator function interacting MHC molecules [29].

Computational analysis of targeted ORF8 RNA device could inhibit the replication of SARS-CoV-2 [30]. Thus, ORF8 is a hotspot for genetic variation in coronaviruses [31]. The clinical effect of deletions in this region appears to be a milder infection with less systemic release of proinflammatory cytokines [32]. An RNA-RNA interaction was reported between SARS-CoV-2 spike and ORF8 genes suggesting ORF8 gene silencing may be required for immune-escape and virus spread [33-34].

Conclusion

Our preliminary study confirmed the presence of many ORF8 truncated mutants in B.1.1.7 lineages of SARS-CoV-2. Although contradictory, we found ORF8 protein L24S mutation in TAA termination codon mutants but no association of L84S mutation was detected. It appeared that such changes neither changed the RNA tertiary structure nor globular interface of ORF8 protein. However, R52I mutation may destroy the histone mimics of ORF8 protein as "ARKS" AA sequence appeared conserved in cellular histones as well as in ORF8 protein of corona virus. None the less, our study indicated a vivid self-destruction process of corona virus due to point mutations eliminating functions of ORF8 protein.

Acknowledgement

The author thanks NCBI database and BLAST search engine for free service worldwide. The free Multiline Software, CLUSTAL-Omega software, Oligoaniline 3.1 software, and SWISS-Model program also acknowledged. AKC is retired professor of Biochemistry.

Competent Interest

The author declares no competing interest

Ethical Issues

The data presented here was computer generated and no animal or human was used.

References

1. Gandhi, L., Misname, D., Rathore, D., Chauhan, P., Bonagura, A., & Venkataraman, M. (2022). Respiratory illness virus infections with special emphasis on COVID-19. *European Journal of Medical Research*, 27(1), 1-21.
2. Lu, G., Wang, Q., & Gao, G. F. (2015). Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-Cove, MERS-Cove, and beyond. *Trends in microbiology*, 23(8), 468-478.
3. Akaishi, T. (2022). Insertion-and-Deletion Mutations between the Genomes of SARS-Cove, SARS-CoV-2, and Bat Coronavirus RaTG13. *Microbiology Spectrum*, e00716-22.
4. Chakraborty A. K. (2020). Coronavirus Nsp2 Protein Homologies to the Bacterial DNA Topoisomerase I and IV Suggest Nsp2 Protein Is a Unique RNA Topoisomerase with Novel Target for Drug and Vaccine Development *Virol & Mycol*, 9: 185.
5. Chakraborty, A. K. (2022) Hyper-variable Spike protein of Omicron corona virus and its differences with Alpha and Delta variants: Prospects of RT-PCR and new vaccine. *J Emerg Dis Virol*. 7 (1):1-13.
6. Chakraborty, A. K. & Chanda, A. (2021). New Biotechnological Exploration on COVID-19 Proteins: Functions, Mutational Profiles and Molecular Targets for Drug Design *SunText Rev Virol*. 2(1), 115.
7. Chakraborty, A. K. (2020). Coronavirus ORF1ab polyprotein associated nsp16 protein is a Rimed Methyltransferase and may methylate 21S mitochondrial rRNA of host cells inhibiting protein synthesis. *Acta Scientific Medical Sciences*, 4 (5): 97-103. DOI: 10.31080/ASMS.2020.04.0629.
8. Hashim, A., Gu, H., Kavain, O., Mori, M., Kwan, M. Y., Chan, W. H., ... & Kavain, N. (2022). SARS-CoV-2 accessory proteins reveal distinct serological signatures in children. *Nature communications*, 13(1), 1-14.
9. Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic acids research*, 16(22), 10881-10890. Yang, Y., Jiang, X.T., Zhang, T. (2014). Evaluation of a Hybrid Approach Using UBLAST and BLASTX for Metagenomic Sequences Annotation of Specific Functional Genes. *PloS One*, 9(10): e110947. Doi: 10.1371/journal.pone.0110947.
10. Waterhouse, A., Bertone, M., Beinart, S., Studer, G., Tauriel, G., Guiney, R., ... & Schweder, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, 46(W1), W296-W303.
11. Elate, S., Martin, H. A., Marczak, R., Fosun, P., Vallejo-Garcia, A., Kumar, G. R., ... & Yukl, S. A. (2022). Novel RT-ddPCR assays for measuring the levels of sub genomic and genomic SARS-CoV-2 transcripts. *Methods*, 201, 15-25.
12. Pereira, F. (2020). Evolutionary dynamics of the SARS-CoV-2 ORF8 accessory gene. *Infection, Genetics and Evolution*, 85, 104525.
13. Ohki, S., Imamura, T., Higashimura, Y., Matsumoto, K., Mori, M. (2021). Similarities and differences in the conformational stability and reversibility of ORF8, an accessory protein of SARS-CoV-2, and its L84S variant. *Biochem Biophys Res Commun*. 563: 92–97. doi: 10.1016/j.bbrc.2021.05.074.
14. Telwatte, S., Martin, H.A., Marczak, R., et al., (2022). Novel RT-ddPCR assays for measuring the levels of sub-genomic and genomic SARS-CoV-2 transcripts.
15. DeRonde, S. Dealing, H., Parker, J., & Chen, J. (2022). Identification of a novel SARS-CoV-2 variant with a truncated protein in ORF8 gene by next generation sequencing. *Scientific reports*, 12(1), 1-8.
16. Takatsuka, H., Fahmi, M., Hanabishi, K., Sakura ani, T., Kubota, Y., & Ito, M. (2022). In silico Analysis of SARSCoV-2 ORF8-Binding Proteins Reveals the Involvement of ORF8 in Acquired-Immune and Innate-Immune Systems. *Frontiers in medicine*, 9.
17. Hassan, S. S., Okahandja, V., Radwan, E. M., Lundstrom, K., Choudhury, P. P., Abd El-Aziz, T. M., ... & Tversky, V. N. (2022). An issue of concern: unique truncated ORF8 protein variants of SARS-CoV-2. *Peer*, 10, e13136.
18. Hassan, S. S., Alkabbani, A. A., Panda, P. K., Ghosh, S., Tartish, D., Choudhury, P. P., ... & Tambuwal, M. M. (2021). A unique view of SARS-CoV-2 through the lens of ORF8 protein. *Computers in biology and medicine*, 133, 104380.
19. Gamage, A. M., Tan, K. S., Chan, W. O., Liu, J., Tan, C. W., Ong, Y. K., ... & Wang, L. F. (2020). Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles. *PLoS pathogens*, 16(12), e1009130.
20. Zhang, Y., Chen, Y., Li, Y., Huang, F., Luo, B., Yuan, Y., ... & Zhang, H. (2021). The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. *Proceedings of the National Academy of Sciences*, 118(23), e2024202118.
21. Chen, X., Zhou, Z., Huang, C., Zhou, Z., Kang, S., Huang, Z., ... & Chen, S. (2021). Crystal Structures of Bat and Human Coronavirus ORF8 Protein Ig-Like Domain Provide Insights Into the Diversity of Immune Responses. *Frontiers in immunology*, 12.
22. Nemudryl, A., Lemuralia, A., Wiegand, T., Nichols, J., Snyder, D. T., Hedges, J. F., ... & Weeniest, B. (2021). SARS-CoV-2 genomic surveillance identifies naturally occurring truncation of ORF7a that limits immune suppression. *Cell reports*, 35(9), 109197.
23. Eucaïne, O. L., Tolani, O., Muhammad, A., Ravichandran, R., Abrash, M. M., Katsuya, S. B., ... & Masaki, M. A. (2021). Translational suppression of SARS-COV-2 ORF8 protein mRNA as a Viable therapeutic target against COVID-19: Computational studies on potential roles of isolated compounds from Cleroterium voluble leaves. *Computers in*

24. Valance, A., Bemuse, A., Álvarez-Bulla, E. R., & Díaz, J. (2021). Structural analysis of SARS-CoV-2 ORF8 protein: pathogenic and therapeutic implications. *Frontiers in genetics*, 12, 693227.
25. Gordon, D. E., Hiatt, J., Boucharde, M., Resell, V. V., Olfert's, S., Barber, H., ... & Durward, M. (2020). Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science*, 370(6521), eabe9403.
26. Arshad, N., Laurent-Rolle, M., Ahmed, W. S., Hsu, J. C. C., Mitchell, S. M., Pawlak, J., ... & Creswell, P. (2022). SARS-CoV-2 accessory proteins ORF7a and ORF3a use distinct mechanisms to downregulate MHC-I surface expression. [Doi:10.1101/2022.05.17.492198](https://doi.org/10.1101/2022.05.17.492198) [Preprint].
27. Goud, V. R., Chakraborty, R., Chakraborty, A., Lavada, K., Patnaik, S., Sharma, S., & Patnaik, S. (2022). A bioinformatic approach of targeting SARS-CoV-2 replication by silencing a conserved alternative reserve of the orf8 gene using host miRNAs. *Computers in biology and medicine*, 145, 105436.
28. Beaudoin-Buskers, G., Arduino, A., Bourassa, C., Mediated, H., Gendron-Lepage, G., Richard, J., ... & Fins, A. (2022). SARS-CoV-2 accessory protein ORF8 decreases antibody-dependent cellular cytotoxicity. *Viruses*, 14(6), 1237.
29. Omori, O. B., Pereira, F., Jenga, S. C., & Manzourolajdad, A. (2022). A Putative long-range RNA-RNA interaction between ORF8 and Spike of SARS-CoV-2. *Plos one* 17(9), e0260331.
30. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410.
31. Badua, C. L. D., Baldor, K. A. T., & Medina, P. M. B. (2021). Genomic and proteomic mutation landscapes of SARS-CoV-2. *Journal of medical virology*, 93(3), 1702-1721.
32. Biswal, M., Diggs, S., Xu, D., Khudaverdyan, N., Lu, J., Fang, J., ... & Song, J. (2021). Two conserved oligomer interfaces of NSP7 and NSP8 underpin the dynamic assembly of SARS-CoV-2 RdRp. *Nucleic acids research*, 49(10), 5956-5966.
33. Chakraborty, A. K. (2020). Multi-Alignment Comparison of Coronavirus Non-Structural Proteins Nsp13-Nsp16 with Ribosomal Proteins and other DNA/RNA Modifying Enzymes Suggested their Roles in the Regulation of Host Protein Synthesis. *Int J Clin Med Info*, 3(1), 7-19.
34. Chakraborty, A. K. (2022). A Method of Identification of SARS-CoV-2 Variant Using NCBI BLAST-2 100% Homology Search with Specific Oligonucleotides Selected at the Deletion Boundaries of S, N, ORF7a, ORF8 and ORF1ab Proteins.
35. Chou, J. M., Tsai, J. L., Hung, J. N., Chen, I. H., Chen, S. T., & Tsai, M. H. (2022). The ORF8 protein of SARS-CoV-2 modulates the spike protein and its implications in viral transmission. *Front Microbiol.* 13, 883597.
36. Fahmi, M., Kitagawa, H., Yesui, G., Kubota, Y., & Ito, M. (2021). The functional classification of ORF8 in SARS-CoV-2 replication, immune evasion, and viral pathogenesis inferred through phylogenetic profiling. *Evolutionary Bioinformatics*, 17, 11769343211003079.
37. Flower, T. G., Buffalo, C. Z., Hooy, R. M., Allaire, M., Ren, X., & Hurley, J. H. (2021). Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proceedings of the National Academy of Sciences*, 118(2), e2021785118.
38. Gong, Y. N., Tsao, K. C., Hsiao, M. J., Huang, C. G., Huang, P. N., Huang, P. W., ... & Shih, S. R. (2020). SARS-CoV-2 genomic surveillance in Taiwan revealed novel ORF8-deletion mutant and clade possibly associated with infections in Middle East. *Emerging microbes & infections*, 9(1), 1457-1466.
39. Harvey, W. T., Carrabelle, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., ... & Robertson, D. L. (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*, 19(7), 409-424.
40. Kee, J., Thulium, S., Renner, D. M., Glassed, K., Palazzolo, K., Zhang, Z., ... & Kerb, E. (2022). SARS-CoV-2 disrupts host epigenetic regulation via histone mimicry. *Nature*, 610(7931), 381-388.
41. Liu, P., Wang, X., Sun, Y., Zhao, H., Cheng, F., Wang, J., ... & Wang, L. (2022). SARS-CoV-2 ORF8 reshapes the ER through forming mixed disulphides with ER oxidoreductases. *Redox Biology*, 54, 102388.
42. Mein Berger, D., Koch, M., Roth, A., Hermes, G., Stemler, J., Cornelly, O. A., ... & Klatt, A. R. (2021). Analysis of IgM, IgA, and IgG isotype antibodies Directed against SARS-CoV-2 spike glycoprotein and ORF8 in the course of COVID-19. *Scientific reports*, 11(1), 1-9.
43. Pangacine, N., Raba ski, L., Grotowski, T., Nowicki, G., Kowalski, M., Widomski, W., ... & Pirc, K. (2021). Expansion of a SARS-CoV-2 Delta variant with an 872 nt deletion encompassing ORF7a, ORF7b and ORF8, Poland, July to August 2021. *Eurosurveillance*, 26(39), 2100902
44. Matsuoka, K., Maharshi, N., Ohno, M., Ode, H., Nakata, Y., Kubota, M., ... & Iwatani, Y. (2022). SARS-CoV-2 accessory protein ORF8 is secreted extracellularly as a glycoprotein homodimer. *Journal of Biological Chemistry*, 298(3).
45. Neches, R. Y., Cyprides, N. C., & Bouzoukis, C. A. (2021). Atypical divergence of SARS-CoV-2 Orf8 from Orf7a within the coronavirus lineage suggests potential stealthy viral strategies in immune evasion. *mBio*, 12(1), e03014-20.
46. Stupor, R. M. (2010). Into the wild: The soybean genome meets its undomesticated relative. *Proceedings of the National Academy of Sciences*, 107(51), 21947-21948.
47. Stupor, R. M. (2010). Into the wild: The soybean genome meets its undomesticated relative. *Proceedings of the National Academy of Sciences*, 107(51), 21947-21948.
48. Su, Y. C., Anderson, D. E., Young, B. E., Linster, M., Zhu, F., Jayakumar, J., ... & Smith, G. J. (2020). Discovery and genomic characterization of a 382-nucleotide deletion in ORF7b and ORF8 during the early evolution of SARS-CoV-2. *MBio*, 11(4), e01610-20.
49. Flower, T. G., Buffalo, C. Z., Hooy, R. M., Allaire, M., Ren, X., & Hurley, J. H. (2021). Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proceedings of the National Academy of Sciences*, 118(2), e2021785118.

-
50. Wang, X., Lam, J. Y., Wong, W. M., Yuen, C. K., Cai, J. P., Au, S. W. N., ... & Yuen, K. Y. (2020). Accurate diagnosis of COVID-19 by a novel immunogenic secreted SARS-CoV-2 orf8 protein. MBio, 11(5), e02431-20.
51. Wang, X., Lam, J. Y., Chen, L., Au, S. W. N., To, K. K., Yuen, K. Y., & Kok, K. H. (2021). Mining of linear B cell epitopes of SARS-CoV-2 ORF8 protein from COVID-19 patients. Emerging microbes & infections, 10(1), 1016-1023.
52. Wu, X., Xia, T., Shin, W. J., Yu, K. M., Jung, W., Herrmann, A., ... & Jung, J. U. (2022). Viral Mimicry of Interleukin-17A by SARS-CoV-2 ORF8. Mbio, 13(2), e00402-22.
53. Yu, T., Ling, Q., Xu, M., Wang, N., Wang, L., Lin, H., ... & Pan, J. A. (2022). ORF8 protein of SARS-CoV-2 reduces male fertility in mice. Journal of Medical Virology, 94(9), 4193-4205.
54. Zinzula, L. (2021). Lost in deletion: The enigmatic ORF8 protein of SARS-CoV-2. Biochemical and biophysical research communications, 538, 116-124.

Copyright: ©2022: Asti Kumar Chakraborty. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.