

1 SARS-CoV-2 spike protein gene variants with N501T and G142D mutation dominated
2 infections in minks in the US

3 Hugh Y. Cai^{1*}, Allison Cai²

4 1. Animal Health Laboratory, University of Guelph, Guelph, Ontario, Canada

5 2. Faculty of Science, University of British Columbia, BC, Canada

6 *Corresponding author: hcai@uoguelph.ca

7 **Abstract**

8 Large number of minks were infected with SARS-CoV-2 virus containing the spike protein
9 Y453F mutation in Europe, causing zoonosis concerns. To evaluate the genetic characteristics of
10 the US and Canadian mink-derived SARS-CoV-2 sequences, we analyzed all animal-derived
11 (977), all Canadian (19,529) and US (173,277) SARS-CoV-2 sequences deposited in GISAID
12 from December 2019 to March 12, 2021, and identified 2 dominant novel variants, the N501T-
13 G142D variant and N501T-G142D-F486L variant, in the US mink-derived SARS-CoV-2
14 sequences. These variants were not found in minks from Canada or other countries. The Y453F
15 mutation was not identified in the mink-derive sequences in the US and Canada. The N501T
16 mutation occurred two months earlier in the human than in the minks in the US, and the novel
17 N501T-G142D variant and N501T-G142D-F486L variant were found in human prior to minks.
18 The result of this study indicates that the novel variants may have evolved during human
19 infection and then transmitted to mink populations in the US.

20 **Keywords**

21 SARS-CoV-2, COVID-19, spike protein, S gene, mutation, variant, mink

22 **Introduction**

23 Genomic surveillance of circulating SARS-CoV-2 variants is critical for epidemiologic
24 tracking and infection control. Rapid global spread of a variant harbouring a mutation located on
25 the SARS-CoV-2 spike protein (S protein) has been reported.^{6,7} In the GISAID database of
26 SARS-CoV-2 genomes as of March 17, 2021, 95.7% (758,684 of 792,388) of the SARS-CoV-2
27 genome has the D614G mutation in the spike protein gene; and 22% (178,186 of 792.388) had

28 N501Y mutation which occurred in the more virulent UK variant (lineage B.1.1.7) with
29 increased transmissibility by 50% and is potentially more virulent;¹ a similar phenomenon was
30 observed with the South African variant lineage B.1.351⁸ and the Brazil variant lineage
31 B.1.525,^{4,9} both have the N501Y mutation. Large numbers of mutants with Y453F mutations
32 were found in infected mink in Denmark, which might have been spread from human to animal
33 and back to human.⁵ We analyzed the sequences of the SARS-CoV-2 S protein gene collected
34 from mink in the USA and Canada, and identified 2 dominant novel variants with 2 and 3
35 mutations each.

36 **Materials & Methods**

37 From the GISAID (Global Initiative on Sharing All Influenza Data, <https://www.gisaid.org>),
38 we downloaded the following SARS-CoV-2 sequences collected from December 2019 to the
39 date of writing (March 12, 2021): all animal-derived sequences (977), all sequences collected in
40 Canada (19,529) and the US (173,277). The SARS-CoV-2 isolate Wuhan-Hu-1 collected in
41 December 19, 2019 (GenBank accession NC045512) was used as a reference for mutation
42 analysis. All nucleotide position labeling in our study was based on alignment with this
43 sequence. SARS-CoV-2 full genome sequences were multiple aligned with the bioinformatic
44 software, Geneious v.11 (Auckland, New Zealand) using “Map to a Reference Assembly”
45 function. The aligned sequences were visually examined to confirm that they were aligned
46 properly. The variants/SNP were identified by the software automatically and verified by visual
47 confirmation. Short fragments (30 nt) containing the novel mutations identified in our study were
48 used as queries to blast search the local databases consisting of the downloaded sequences to
49 determine the incidence of the novel mutations using the “BLAST” function of the Geneious
50 software. The global incidence of “N501Y” was determined using the “search” function with
51 substitution “N501Y” in GISAID. The global incidence of “N501T” was not determined since
52 the query of substitution “N501T” was not available in GISAID.

53 **Results and Discussion**

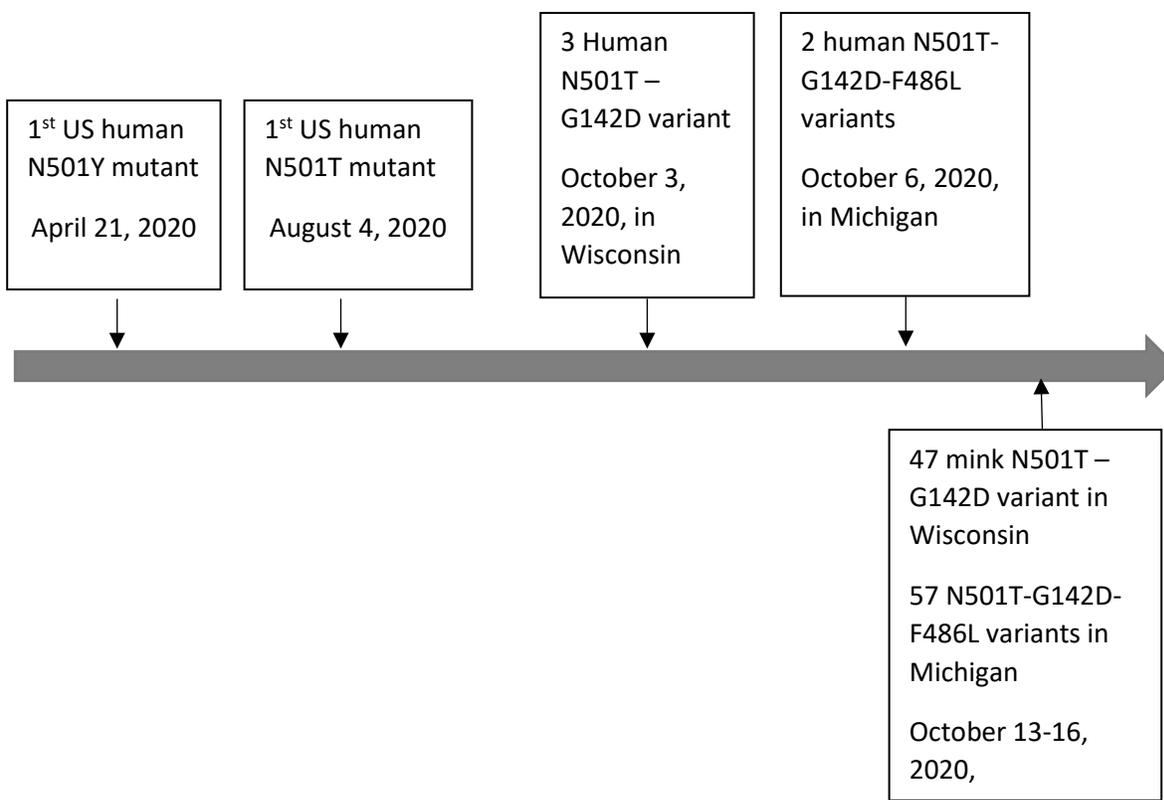
54 The N501Y mutation that was common to UK, South African, and Brazil variants, and the
55 Y453F mutation that widely prevailed in COVID-19 positive mink in Denmark and the
56 Netherlands were absent in all the US and Canadian mink SARS-CoV-2 sequences examined.
57 Interestingly, an N501T mutation was found in almost all US mink SARS-CoV-2 sequences (100

58 of 101; 99%); furthermore, all N501T mutant sequences, except 2 of poor quality, had the
59 G142D mutation (Supplementary Table 1). Based on GISAID records, all US mink SARS-CoV-
60 2 sequences were collected from Michigan and Wisconsin in October, 2020. We found that,
61 except 1 sequences with poor quality, all (54) Michigan mink sequences had the F486L mutation
62 in addition to the N501T and G142D mutations (Supplementary Table 1). Hereafter, we
63 designated the sequences with N501T and G142D, but without F486L, mutations as the N501T-
64 G142D variant; and the ones with all 3 novel mutations as the N501T-G142D-F486L variant. In
65 addition to the above 3 mutations, G614D, the mutation that dominated the SARS-CoV-2
66 sequences, existed in all US and Canadian mink-derived sequences (Supplementary Table 1,
67 Supplementary Table 2). No N501T-G142D nor N501T-G142D-F486L variants were found in
68 mink from Canada or other countries. Outside the US, among 335 Netherlands mink-derived
69 SARS-CoV-2 S protein gene sequences, 5 collected from April to June 2020 had the N501T
70 mutation without G142D or F486L mutations (Supplementary Table 3), which is consistent with
71 a previous report on N501T on Netherlands mink-derived sequences.² We also found F486L
72 mutation in 60% (204 of 335) Netherlands mink-derived SARS-CoV-2 S protein gene
73 sequences, none of which with N501T-G142D variants. The F486L mutation was unique to
74 Netherlands mink-derived sequences causing by nucleotide mutation T23018C, different from
75 the F486L mutation found in the US mink-derived sequences due to nucleotide T23020G
76 mutation. Among 454 Denmark mink SARS-CoV-2 sequences, 3 had N501T mutations, none
77 had G142D or F486L mutations (Supplementary Table 3).

78 In order to determine the origin of the 2 novel variants, we analyzed all human-derived
79 SARS-CoV-2 sequences collected from the US and Canada. Three N501T-G142D variant and 2
80 N501T-G142D-F486L variant sequences were identified from the sequences collected in early
81 October 2020 from Wisconsin and Michigan, respectively. Furthermore, we identified N501T
82 mutants without G142D and F486L in US human-derived sequences collected as early as August
83 2020, which was earlier than the mink-derived sequences collected in October, 2020
84 (Supplementary Table 4, Fig. 1). In Canada, the N501T mutation only existed in human-derived
85 sequences (Supplementary Table 2). The N501T mutation was described in 2 Italian human-
86 derived sequences collected in August 2020.³ We therefore speculate that the N501T mutation
87 may have evolved during human infection and was then transmitted to mink populations in the
88 US, and it is less likely that the mutations occurred in mink population after infected with human

89 strains of the virus. The fact that the two novel variants were found in almost all infected minks
90 in the US in short period of times within days also support this speculation.

91 To evaluate the transmissibility of N501T mutants, the incidence of N501T and N501Y
92 mutants was enumerated from US and Canadian GISAID SARS-CoV-2 sequences collected up
93 to March 12, 2021: 2.4% (4,097 of 173,277) and 0.77% (1,339 of 173,277) of US sequences had
94 the N501Y and N501T maturation, respectively; and 1.6% (315/19,529) and 0.06% (12/19,529)
95 of Canadian sequences had the N501Y and the N501T mutation, respectively (Supplementary
96 Table 5). Comparatively, the incidence of N501Y is much lower in the US and Canada than
97 globally. Our analysis results indicate that the N501T mutants had weaker transmissibility
98 compared to the N501Y mutants that include UK B.1.17, South Africa B.1.351, and Brazil P.1.
99 However, it cannot be ruled out that the different time of introduction of N501Y and N501T
100 (April 2020 vs August 2020) may have affected the occurrences of the 2 mutants (Fig. 1).



101 Fig. 1. Emerging timeline of human-derived (above the timeline) and mink-derived (below the
102 timeline) SARS-CoV-2 spike protein N501T-G142D and N501T-G142D-F486L variants in the
103 US. Timeline scale is not in proportion.

104

105 It has been described that the mutations at S protein N501 can significantly increase the
106 transmissibility, and likely virulence of SARS-CoV-2; the UK variant N501Y is an example. It
107 was predicted that a single N501T mutation might significantly enhance the binding affinity
108 between 2019-nCoV RBD and human ACE2.¹⁰ Our analysis revealed that the N501T and
109 G142DL mutations occurred in 99% of mink-derived sequences collected in the US; all, except
110 one with poor quality, mink sequences collected from Michigan also contained the F486L
111 mutation. The large number of new variants occurring in the mink population warrants further
112 study on how these changes affected their interaction with the ACE2 receptor, and thereby the
113 transmissibility, virulence, and immunogenicity in humans and mink. It is important to monitor
114 the emerging new variants and determine their impact on human and animal health.

115

116 **Acknowledgement**

117 We gratefully acknowledge the authors, originating and submitting laboratories of the sequences
118 from GISAID's hCoV-19 Database on which this research is based. See supplementary
119 acknowledgement for details.

120

121 **Declaration of conflicting interests**

122 The authors declared no potential conflicts of interest with respect to the research, authorship,
123 and/or publication of this article.

124

125 **Funding Support:**

126 The authors received no financial support for the research, authorship, and/or publication of this
127 article.

128 **Supplementary material**

129 Supplementary material for this article is available online.

130 **References**

131

- 132 1. ECDC (21 January 2021). "Risk related to the spread of new SARS-CoV-2 variants of
133 concern in the EU/EEA - first update". European Centre for Disease Prevention and
134 Control. Retrieved 2 February 2021.
- 135 2. Elasad A, Fawzy M, Basiouni S, and ShehataAA. Mutational spectra of SARS-CoV-2
136 isolated from animals PeerJ. 8: e10609. Published online 2020 Dec 18. doi:
137 [10.7717/peerj.10609](https://doi.org/10.7717/peerj.10609)
- 138 3. Fiorentini S et al. First detection of SARS-CoV-2 spike protein N501 mutation in Italy in
139 August, 2020. Lancet Infect Dis 2021 Published Online, January 12, 2021,
140 [https://doi.org/10.1016/S1473-3099\(21\)00007-4](https://doi.org/10.1016/S1473-3099(21)00007-4) (Retrieved March 12, 2021)
- 141 4. Kupferschmidt K. New mutations raise specter of 'immune escape. Science. 371 (6527):
142 329–330. doi:10.1126/science.371.6527.329. PMID 33479129. Retrieved 25 January
143 2021.
- 144 5. Munnink BBO, et al. Transmission of SARS-CoV-2 on mink farms between humans and
145 mink and back to humans. Science. 2021; 371:172-177 DOI: 10.1126/science.abe5901
- 146 6. Na W, Moon H and Song DA. Comprehensive review of SARS-CoV-2 genetic mutations
147 and lessons from animal coronavirus recombination in one health perspective. J
148 Microbiol. 2021. 59, 332–340 <https://doi.org/10.1007/s12275-021-0660-4>
- 149 7. Phan T. Genetic diversity and evolution of SARS-CoV-2. Infect Genet Evol. 2020;
150 81:104260
- 151 8. Tegally H et al. Emergence and rapid spread of a new severe acute respiratory
152 syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in
153 South Africa medRxiv 2020.12.21.20248640; doi:
154 <https://doi.org/10.1101/2020.12.21.20248640>
- 155 9. Voloch, CM et al. Genomic characterization of a novel SARS-CoV-2 lineage from Rio
156 de Janeiro, Brazil . Retrieved 15 January 2021. doi:10.1101/2020.12.23.20248598 – via
157 medRxiv.

- 158 10. Wan Y, et al. Receptor recognition by the novel coronavirus from Wuhan: an analysis
159 based on decade-long structural studies of SARS coronavirus. J Virol 2020; 94:e00127-
160 20. <https://doi.org/10.1128/JVI.00127-20>.