

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<sup>1</sup>, Allison Cai<sup>2</sup>

4 Animal Health Laboratory, University of Guelph, Guelph, Ontario, Canada (H. Cai)

5 Faculty of Science, University of British Columbia, BC, Canada (A. Cai)

6

7 \*Corresponding author: Hugh Cai, Animal Health Laboratory, University of Guelph, 419

8 Gordon Street, Bldg 89, Guelph, Ontario, Canada. N1G 2W1. hcai@uoguelph.ca

9

10 Running title: SARS-CoV-2 N501T-G142D variants dominated infections in minks in the US

## 11 **Abstract**

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

## 24 **Keywords**

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

26

27 Genomic surveillance of circulating SARS-CoV-2 variants is critical for epidemiologic  
28 tracking and infection control. Rapid global spread of a variant harbouring a mutation located on  
29 the SARS-CoV-2 spike protein (S protein) has been reported.<sup>6,7</sup> In the GISAID database of  
30 SARS-CoV-2 genomes as of March 17, 2021, 95.7% (758,684 of 792,388) of the SARS-CoV-2  
31 genome has the D614G mutation in the spike protein gene; and 22% (178,186 of 792.388) had  
32 N501Y mutation which occurred in the more virulent UK variant (lineage B.1.1.7) with  
33 increased transmissibility by 50% and is potentially more virulent;<sup>1</sup> a similar phenomenon was  
34 observed with the South African variant lineage B.1.351<sup>8</sup> and the Brazil variant lineage  
35 B.1.525,<sup>4,9</sup> both have the N501Y mutation. Large numbers of mutants with Y453F mutations  
36 were found in infected mink in Denmark, which might have been spread from human to animal  
37 and back to human.<sup>5</sup> We analyzed the sequences of the SARS-CoV-2 S protein gene collected  
38 from mink in the USA and Canada, and identified 2 dominant novel variants with 2 and 3  
39 mutations each.

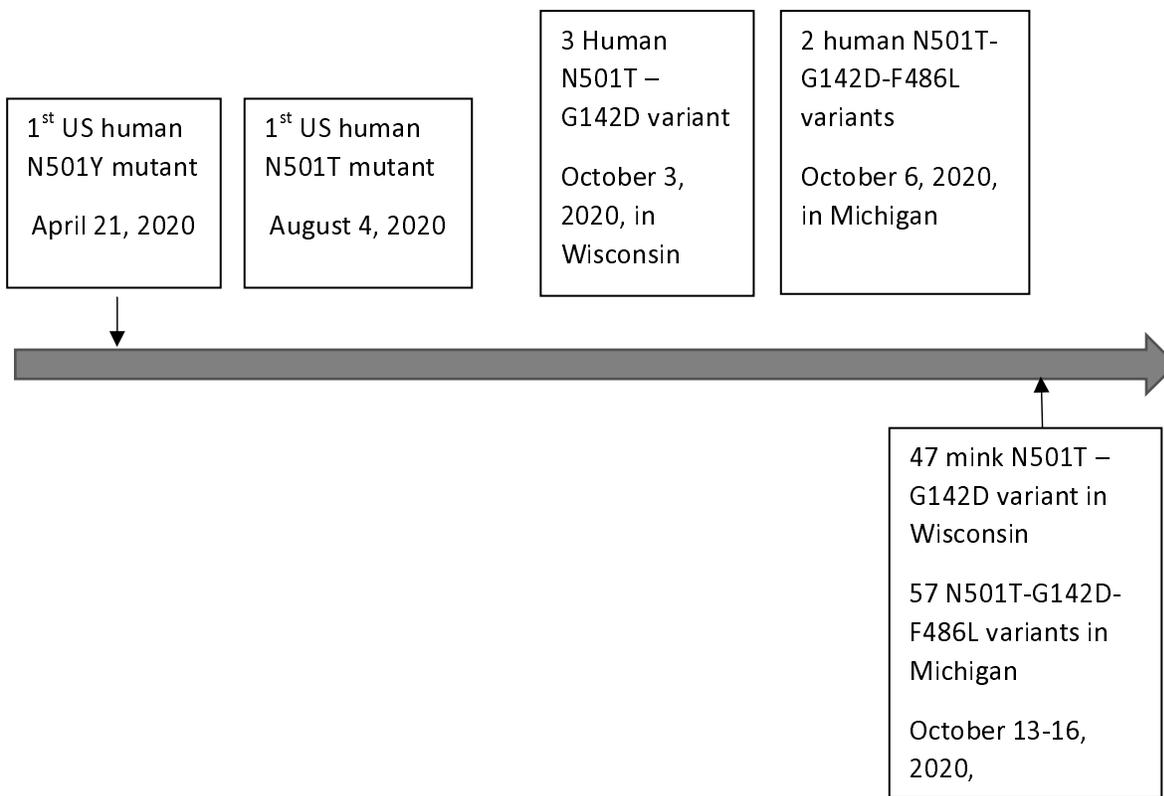
40 From the GISAID (Global Initiative on Sharing All Influenza Data, <https://www.gisaid.org>),  
41 we downloaded the following SARS-CoV-2 sequences collected from December 2019 to the  
42 date of writing (March 12, 2021): all animal-derived sequences (977) including those from *Canis*  
43 *lupus familiaris*, *Chlorocebus sabaeus*, *Felis catus*, *Manis javanica*, *Manis pentadactyla*, *Mus*  
44 *musculus*, *Mustela lutreola*, *Neovison vison*, *Panthera leo*, *Panthera tigris jacksoni*, *Rhinolophus*  
45 *affinis*, *Rhinolophus malayanus* and *Rhinolophus shameli*; all sequences collected in Canada  
46 (19,529) and the US (173,277). The SARS-CoV-2 isolate Wuhan-Hu-1 collected in December  
47 19, 2019 (GenBank accession NC045512) was used as a reference for mutation analysis. All  
48 nucleotide position labeling in our study was based on alignment with this sequence. SARS-  
49 CoV-2 full genome sequences were multiple aligned with the bioinformatic software, Geneious  
50 v.11 (Auckland, New Zealand) using “Geneious Multiple Alignment” and “Map to a Reference  
51 Assembly” function. The aligned sequences were visually examined to confirm that they were  
52 aligned properly. The variants/SNP were identified by the software automatically and verified by  
53 visual confirmation. Short fragments (30 nt) containing the novel mutations identified in our  
54 study were used as queries to blast search the local databases consisting of the downloaded  
55 sequences to determine the incidence of the novel mutations using the “BLAST” function of the  
56 Geneious software. The global incidence of “N501Y” was determined using the “search”  
57 function with substitution “N501Y” in GISAID. The global incidence of “N501T” was not  
58 determined since the query of substitution “N501T” was not available in GISAID.

59 The N501Y mutation that was common to UK, South African, and Brazil variants, and the  
60 Y453F mutation that widely prevailed in COVID-19 positive mink in Denmark and the  
61 Netherlands were absent in all the US and Canadian mink SARS-CoV-2 sequences examined.  
62 Interestingly, an N501T mutation was found in almost all US mink SARS-CoV-2 sequences (100  
63 of 101; 99%); furthermore, all N501T mutant sequences, except 2 of poor quality, had the  
64 G142D mutation (Supplementary Table 1). Based on GISAID records, all US mink SARS-CoV-  
65 2 sequences were collected from Michigan and Wisconsin in October, 2020. We found that,  
66 except 1 sequences with poor quality, all (54) Michigan mink sequences had the F486L  
67 mutation, and all (55) Michigan mink sequences had the Y144 deletion, in addition to the N501T  
68 and G142D mutations (Supplementary Table 1). Hereafter, we designated the sequences with  
69 N501T and G142D, but without F486L, mutations as the N501T-G142D variant (WI mink  
70 variant); and the ones with all 3 novel mutations as the N501T-G142D-F486L variant (MI mink

71 variant). In addition to the above 3 mutations and 1 deletion in Spike protein, G614D, the  
72 mutation that dominated the SARS-CoV-2 sequences, existed in all US and Canadian mink-  
73 derived sequences (Supplementary Table 1, Supplementary Table 2). No N501T-G142D nor  
74 N501T-G142D-F486L variants were found in mink from Canada or other countries.

75 Outside the US, among 335 Netherlands mink-derived SARS-CoV-2 S protein gene  
76 sequences, 5 collected from April to June 2020 had the N501T mutation without G142D or  
77 F486L mutations (Supplementary Table 3), which is consistent with a previous report on N501T  
78 on Netherlands mink-derived sequences.<sup>2</sup> We also found F486L mutation in 60% (204 of 335)  
79 Netherlands mink-derived SARS-CoV-2 S protein gene sequences, none of which with N501T-  
80 G142D variants. The F486L mutation was unique to Netherlands mink-derived sequences  
81 causing by nucleotide mutation T23018C, different from the F486L mutation found in the US  
82 mink-derived sequences due to nucleotide T23020G mutation. Among 454 Denmark mink  
83 SARS-CoV-2 sequences, 3 had N501T mutations as described earlier, none had G142D or  
84 F486L mutations (Supplementary Table 3).

85 In order to determine the origin of the 2 novel variants, we analyzed all human-derived  
86 SARS-CoV-2 sequences collected from the US and Canada. Three WI N501T-G142D variant  
87 and 2 MI N501T-G142D-F486L variant sequences were identified from the US human-derived  
88 sequences collected from Wisconsin on October 3, 2020 and Michigan on October 6, 2020, and  
89 they had all the mutations as those described above for the mink sequences except lacking of the  
90 Y144 deletion in S gene. Furthermore, we identified N501T mutants without G142D and F486L  
91 in US human-derived sequences collected as early as August 2020, which was earlier than the  
92 mink-derived sequences collected in October, 2020 (Supplementary Table 4, Fig. 1). In Canada,  
93 the N501T mutation only existed in human-derived sequences (Supplementary Table 2). The  
94 N501T mutation was described in 2 Italian human-derived sequences collected in August 2020.<sup>3</sup>  
95 We therefore speculate that the N501T mutation may have evolved during human infection and  
96 was then transmitted to mink populations in the US, and it is less likely that the mutations  
97 occurred in mink population after infected with human strains of the virus. The fact that the two  
98 novel variants were found in almost all infected minks in the US in short period of times within  
99 days also support this speculation.



100  
101  
102

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

106

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

117

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

128

### 129 **Acknowledgement**

130 We gratefully acknowledge the authors, originating and submitting laboratories of the sequences  
131 from GISAID's hCov-19 Database on which this research is based. See supplementary  
132 acknowledgement for details.

133

### 134 **Declaration of conflicting interests**

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

137

### 138 **Funding Support:**

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

### 141 **Supplementary material**

142 Supplementary material for this article is available online.

143

144 **References**

145

- 146 1. ECDC (21 January 2021). "Risk related to the spread of new SARS-CoV-2 variants of  
147 concern in the EU/EEA - first update". European Centre for Disease Prevention and  
148 Control. Retrieved 2 February 2021.
- 149 2. Elasad A, Fawzy M, Basiouni S, and ShehataAA. Mutational spectra of SARS-CoV-2  
150 isolated from animals PeerJ. 8: e10609. Published online 2020 Dec 18. doi:  
151 10.7717/peerj.10609
- 152 3. Fiorentini S et al. First detection of SARS-CoV-2 spike protein N501 mutation in Italy in  
153 August, 2020. Lancet Infect Dis 2021 Published Online, January 12, 2021,  
154 [https://doi.org/10.1016/S1473-3099\(21\)00007-4](https://doi.org/10.1016/S1473-3099(21)00007-4) (Retrieved March 12, 2021)
- 155 4. Kupferschmidt K. New mutations raise specter of 'immune escape. Science. 371 (6527):  
156 329–330. doi:10.1126/science.371.6527.329. PMID 33479129. Retrieved 25 January  
157 2021.
- 158 5. Munnink BBO, et al. Transmission of SARS-CoV-2 on mink farms between humans and  
159 mink and back to humans. Science. 2021; 371:172-177 DOI: 10.1126/science.abe5901
- 160 6. Na W, Moon H and Song DA. Comprehensive review of SARS-CoV-2 genetic mutations  
161 and lessons from animal coronavirus recombination in one health perspective. J  
162 Microbiol. 2021. 59, 332–340 <https://doi.org/10.1007/s12275-021-0660-4>
- 163 7. Phan T. Genetic diversity and evolution of SARS-CoV-2. Infect Genet Evol. 2020;  
164 81:104260
- 165 8. Tegally H et al. Emergence and rapid spread of a new severe acute respiratory  
166 syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in  
167 South Africa medRxiv 2020.12.21.20248640; doi:  
168 <https://doi.org/10.1101/2020.12.21.20248640>
- 169 9. Voloch, CM et al. Genomic characterization of a novel SARS-CoV-2 lineage from Rio  
170 de Janeiro, Brazil . Retrieved 15 January 2021. doi:10.1101/2020.12.23.20248598 – via  
171 medRxiv.
- 172 10. Wan Y, et al. Receptor recognition by the novel coronavirus from Wuhan: an analysis  
173 based on decade-long structural studies of SARS coronavirus. J Virol 2020; 94:e00127-  
174 20. <https://doi.org/10.1128/JVI.00127-20>.

175

176