

1 **Mutational cascade of SARS-CoV-2 leading to evolution and emergence of omicron** 2 **variant**

3 **Kanika Bansal^{1*}, Sanjeet Kumar^{2*}**

4 1 CSIR-Institute of Microbial Technology, Chandigarh, India

5 2 Gangadhar Meher University, Sambalpur, Odisha, India

6 * Corresponding author

7 Address for correspondence:

8 kanikabansal@imtech.res.in (K. Bansal); ksanjeet.ibab@imtech.res.in

9 **Abstract**

10 **Background:** Emergence of new variant of SARS-CoV-2, namely omicron, has posed a
11 global concern because of its high rate of transmissibility and mutations in its genome.
12 Researchers worldwide are trying to understand the evolution and emergence of such variants
13 to understand the mutational cascade events.

14 **Methods:** We have considered all omicron genomes (n = 302 genomes) available till 2nd
15 December 2021 in the public repository of GISAID along with representatives of variants of
16 concern (VOC), i.e., alpha, beta, gamma, delta, and omicron; variant of interest (VOI) mu
17 and lambda; and variant under monitoring (VUM). Whole genome-based phylogeny and
18 mutational analysis were performed to understand the evolution of SARS CoV-2 leading to
19 emergence of omicron variant.

20 **Results:** Whole genome-based phylogeny depicted two phylogroups (PG-I and PG-II)
21 forming variant specific clades except for gamma and VUM GH. Mutational analysis
22 detected 18,261 mutations in the omicron variant, majority of which were non-synonymous
23 mutations in spike (A67, T547K, D614G, H655Y, N679K, P681H, D796Y, N856K, Q954H),
24 followed by RNA dependent RNA polymerase (rdrp) (A1892T, I189V, P314L, K38R, T492I,
25 V57V), ORF6 (M19M) and nucleocapsid protein (RG203KR).

26 **Conclusion:** Delta and omicron have evolutionary diverged into distinct phylogroups and do
27 not share a common ancestry. While, omicron shares common ancestry with VOI lambda and
28 its evolution is mainly derived by the non-synonymous mutations.

29 **Abbreviations:**

30 VOC, variant of concern; VOI, variant of interest; VUM, variant under monitoring; NSP,
31 non-structural protein; UTR, untranslated region; rdrp, RNA dependent RNA polymerase.

32 **Keywords:** SARS-CoV-2, COVID-19, genome-wide, evolution, deadly variants, VOC, VOI,
33 VUM, SNP, mutation, non-synonymous, silent mutation, spike, RNA dependent RNA
34 polymerase, NSP, UTR

35 **Introduction**

36 Currently, the world is witnessing a resurgence of COVID-19 cases due to the new omicron
37 variant belonging to B.1.1.529. Omicron was first reported in South Africa on 24th November
38 2021 from the specimen collected on 9th November
39 2021([https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-\(b.1.1.529\)-
40 technical-brief-and-priority-actions-for-member-states](https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-(b.1.1.529)-technical-brief-and-priority-actions-for-member-states)). On 26th November 2021, World
41 Health Organisation (WHO) assigned omicron to the ‘variant of concern’ (VOC) category
42 due to its ability to poses a higher risk of reinfection as compared to previously reported
43 variants ([https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-
44 sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern); [https://www.who.int/news/item/28-11-2021-update-on-
45 omicron](https://www.who.int/news/item/28-11-2021-update-on-omicron)). According to the 1st December 2021 update, omicron is reported in at least 23
46 countries from five out of six WHO regions, with most cases in Africa and Europe
47 ([https://www.cnbc.com/2021/12/01/who-says-omicron-has-been-found-in-23-countries-
48 across-the-world.html](https://www.cnbc.com/2021/12/01/who-says-omicron-has-been-found-in-23-countries-across-the-world.html)).

49 There is a lot of uncertainty surrounding the omicron variant. For its risk assessment,
50 scientists and researchers are investigating the intensity of its spread, extent of its infection,
51 effectiveness of detection methods, therapeutics, and vaccine efficacy (Knoll & Wonodi,
52 2021; Lipsitch & Dean, 2020; Pegu et al., 2021). The onset of omicron is reported with mild
53 diseases suggests its low or mild severity than its previous counterparts like delta (Ewen
54 Callaway, 2021; E. Callaway & Ledford, 2021). It is known to have a very high mutation rate
55 with more than 30 mutational changes in its spike protein (Ewen Callaway, 2021)
56 ([https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-\(b.1.1.529\)-
57 technical-brief-and-priority-actions-for-member-states](https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-(b.1.1.529)-technical-brief-and-priority-actions-for-member-states))

58 Globally, high risk of reinfection with omicron variant, low vaccine and testing coverage are
59 ideal for mutations resulting in the emergence of new variants of SARS-CoV-2 (Pulliam et

60 al., 2021). Since COVID-19 inception, researchers have been trying to investigate its origin
61 and evolution (Bansal, Kumar, & Patil, 2021; Singh & Soojin, 2021; Tang et al., 2020). We
62 are currently witnessing a global molecular arms race between SARS-CoV-2 and its
63 preventive therapeutics based on diverse regimes such as DNA, RNA, protein or inactivated
64 whole-virion, etc. (Andreadakis et al., 2020; Corey, Mascola, Fauci, & Collins, 2020;
65 Sharma, Sultan, Ding, & Triggler, 2020). This global crisis can be addressed by a very rapid
66 immunization program worldwide. Moreover, the real-time monitoring of evolutionary
67 cascade of SARS-CoV-2 leading to lethal variants is utmost. Earlier investigation of several
68 VOC and VOI suggests some of the crucial mutations for viral survival and high infectivity
69 in humans (Boehm et al., 2021; Kumar & Bansal, 2021; Schmidt et al., 2021). However,
70 mutations giving rise to omicron and intra-omicron genomic diversity are not yet analyzed at
71 a population level.

72 In the present study, we aim to look for the mutational profile of deadly and under-
73 monitoring variants reported till now to understand the emergence of a heavily mutated
74 variant named omicron. Interestingly, whole genome-based phylogeny suggests two major
75 phylogroups PG-I and PG-II. Further, mutational analysis depicted the key role of non-
76 synonymous mutations in the evolution of lethal variants. Such genomic insights into the
77 nature of mutational cascade are the need of the hour.

78 **Results**

79 **Phylogenomics suggests common ancestry of omicron and lambda variants**

80 Whole genome-based phylogeny (n = 478 genomes) representing VOC (alpha, beta, gamma,
81 delta, and omicron), VOI (mu and lambda), VUM depicts two major phylogroups PG-I and
82 PG-II (figure 1, supplementary figure 1 and supplementary table 1). Here, the reference strain
83 of SARS-CoV-2 (Wuhan-Hu-1, NC_045512.2) is taken as an outgroup. PG-I has VOC:
84 gamma, beta, and delta; VOI: mu and VUM: GH. PG-II includes VOC: alpha, omicron and
85 VOI: lambda. However, seven strains of VOC gamma are in PG-II, two of which are basal to
86 PG-II forming its outgroup (EPI_ISL_3218258, EPI_ISL_2454057), and five
87 (EPI_ISL_2220217, EPI_ISL_2216321, EPI_ISL_2223074, EPI_ISL_2224081, and
88 EPI_ISL_2224090) are more related to alpha variant. EPI_ISL_3160245 one of the VUM
89 strain is distant from the main clade of VUM.

90 Interestingly, two deadly VOCs, delta and omicron, belong to different phylogroups.
91 Phylogeny depicted that omicron shares a common ancestry with VOI lambda represented by

92 a black asterisk in figure 1. Interestingly, three isolated from Italy (EPI_ISL_6854346,
93 EPI_ISL_6854347, and EPI_ISL_6854348) form a diversified sub-lineage among the
94 omicron population. Additionally, EPI_ISL_6886594 from Germany is a diversified omicron
95 strain.

96 **Very high non-synonymous mutations give rise to omicron**

97 To further understand the evolution and emergence of omicron, we have performed a
98 mutational analysis with respect to the reference genome of SARS-CoV-2 (NC_045512.2)
99 (figure 2). Total mutations detected in the dataset were 24,189, and omicron genomes
100 constituted 18,261 (supplementary table 2). Interestingly, >97% (n = 17,703 mutations) of the
101 mutations in omicron were in the coding region, and remaining 558 were detected in the
102 extragenic region of the genome. Amongst the coding gene mutations, 2,965 were indels
103 while 14,738 were SNPs constituting non-synonymous (n = 11,995 mutations) and
104 synonymous mutations (n = 2,743 mutations).

105 Interestingly, mutational events are highly skewed towards the spike protein, which
106 constitutes ~60% (n = 10,658) of the total mutations in the coding genomic region (n =
107 17,703) (figure 3). The majority of spike protein mutations encompass A67, T547K, D614G,
108 H655Y, N679K, P681H, D796Y, N856K, Q954H, which are reported in all the omicron
109 genomes analyzed (table 3). Count of mutations in the spike was followed by RNA
110 dependent RNA polymerase (rdp) (n = 4,142) constituting A1892T, I189V, P314L, K38R,
111 T492I, V57V in all omicron genomes analyzed (figure 3 and table 3). Remaining 2903
112 mutations were detected in rest of the coding genomic region (table 2, 3, and supplementary
113 table 1), where M19M in ORF6, and RG203KR in nucleocapsid protein are amongst the most
114 prevalent mutations in omicron (figure 3).

115 **Low intra-sequence diversity amongst omicron variant**

116 Intra-strain diversity among the omicron variant strains reported worldwide will be crucial in
117 understanding the genome dynamics and rapid evolution of SARS-JCoV-2. We performed
118 the mutational analysis on the current dataset using omicron (OL677199) isolated from
119 Canada on 23rd November 2021 as the reference genome. Most of the strains (n = 298),
120 irrespective of their geographic origin, had less than ten mutations depicting low intra-strain
121 diversity among omicron strains. However, four of the isolates two from Europe (Italy)
122 (EPI_ISL_6854347 (n = 23 mutations) and EPI_ISL_6854346 (n = 14 mutations) and two
123 from South Africa (EPI_ISL_6699742 (n = 12 mutations) and EPI_ISL_6774091 (n = 11

124 mutations) were most diversified among the omicron genomes. While omicron had >55
125 mutations when compared with other VOCs and VOIs.

126 **Methods**

127 **Identification and procurement of SARS-CoV-2 genome from the public repository**

128 We have considered all the available genomes of omicron variant available in public domain
129 until 6 pm Indian Standard Time (IST) on 2nd December 2021 from GISAID (n = 302
130 genomes). A total of 25 strains from each variant of concern, namely alpha (B.1.1.7), beta
131 (B.1.351), gamma (P.1) and delta (B.1.617.2) and variant of interest, namely lambda (C.37)
132 and mu (B.1.621). We have also considered 25 strains from variant under monitoring, namely
133 GH (B.1.640). These all strains are from their respective earlier reports in the public domain.
134 A detailed list of all the strains used in the study is provided in the supplementary information
135 (supplementary table 1).

136 **Phylogenetic analysis**

137 A total of 477 high-quality genomes, including the major variants spread across the globe
138 were taken into consideration. Multiple sequence alignment was performed for all the
139 genomes using MAFFT v7.467 (Nakamura, Yamada, Tomii, & Katoh, 2018) followed by
140 phylogenetic tree construction using fasttree v2.1.8 with double precision (Price, Dehal, &
141 Arkin, 2010) with gamma time reversal method. Visualization of the obtained phylogenetic
142 tree was performed using iTol v6 (Letunic & Bork, 2019). Different variants were marked in
143 accordance with different colors as mentioned in the legends.

144 **Mutational analysis**

145 Mutational analysis of all the strains (n=477) in the study was performed with two different
146 reference genomes. First with NC_045512.2 (Wuhan-Hu-1) strain (reference SARS CoV-2
147 strain) and another with first reported strain of omicron variant (OL677199.1)
148 (<https://www.ncbi.nlm.nih.gov/nuccore/OL677199>) using nucmer v3.1 (Delcher, Phillippy,
149 Carlton, & Salzberg, 2002). We have used a well-documented R script described earlier
150 (Mercatelli & Giorgi, 2020). Here, we have used gff3 annotation and reference genome file to
151 extract genomic coordinate of SARS-CoV-2 proteins. R library package seqinr ([https://cran.r-](https://cran.r-project.org/web/packages/seqinr/index.html)
152 [project.org/web/packages/seqinr/index.html](https://cran.r-project.org/web/packages/seqinr/index.html)) and biostring package
153 (<https://bioconductor.org/packages/release/bioc/html/Biostrings.html>) of bioconductor was
154 implemented to obtain the list of all the mutational events. Frequency and rate of mutation
155 were calculated with respect to two different references (Reference SARS CoV-2 strain:

156 NC_045512.2) (https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2) and omicron
157 (OL677199.1) (<https://www.ncbi.nlm.nih.gov/nuccore/OL677199>) separately.

158 **Funding Information**

159 Nil

160 **Author contribution statement**

161 Kanika Bansal performed data curation, analysis and writing of manuscript with inputs from
162 Sanjeet Kumar.

163 **Declaration of Competing Interest**

164 The author declares no competing interest.

165 **Acknowledgement**

166 Authors acknowledge the support and motivation from Dr. Prabhu B.Patil – CSIR-Institute of
167 Microbial Technology, Chandigarh. We are also thankful to Dr. Santosh Kumar Sethi for his
168 kind support during the process of study. We also acknowledge GISAID initiative for
169 extensive curation and availability of genomic resource in public domain.

170 **References**

- 171 Andreadakis, Z., Kumar, A., Román, R. G., Tollefsen, S., Saville, M., & Mayhew, S. (2020). The COVID-
172 19 vaccine development landscape. *Nature reviews. Drug discovery*, 19(5), 305-306.
- 173 Bansal, K., Kumar, S., & Patil, P. B. (2021). Codon usage pattern reveals SARS-CoV-2 as a
174 monomorphic pathogen of hybrid origin with role of silent mutations in rapid evolutionary
175 success. *bioRxiv*, 2020.2010.2012.335521. doi:10.1101/2020.10.12.335521
- 176 Boehm, E., Kronig, I., Neher, R. A., Eckerle, I., Vetter, P., & Kaiser, L. (2021). Novel SARS-CoV-2
177 variants: the pandemics within the pandemic. *Clinical Microbiology and Infection*.
- 178 Callaway, E. (2021). Heavily mutated coronavirus variant puts scientists on alert. *Nature*.
- 179 Callaway, E., & Ledford, H. (2021). How bad is Omicron? What scientists know so far. *Nature*.
180 doi:10.1038/d41586-021-03614-z
- 181 Corey, L., Mascola, J. R., Fauci, A. S., & Collins, F. S. (2020). A strategic approach to COVID-19 vaccine
182 R&D. *Science*, 368(6494), 948-950.
- 183 Delcher, A. L., Phillippy, A., Carlton, J., & Salzberg, S. L. (2002). Fast algorithms for large-scale
184 genome alignment and comparison. *Nucleic acids research*, 30(11), 2478-2483.
- 185 Knoll, M. D., & Wonodi, C. (2021). Oxford–AstraZeneca COVID-19 vaccine efficacy. *The Lancet*,
186 397(10269), 72-74.
- 187 Kumar, S., & Bansal, K. J. V. r. (2021). Cross-sectional genomic perspective of epidemic waves of
188 SARS-CoV-2: a pan India study. 198642.
- 189 Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new
190 developments. *Nucleic acids research*, 47(W1), W256-W259.
- 191 Lipsitch, M., & Dean, N. E. (2020). Understanding COVID-19 vaccine efficacy. *Science*, 370(6518), 763-
192 765.
- 193 Mercatelli, D., & Giorgi, F. M. (2020). Geographic and genomic distribution of SARS-CoV-2 mutations.
194 *Frontiers in microbiology*, 11, 1800.

- 195 Nakamura, T., Yamada, K. D., Tomii, K., & Katoh, K. (2018). Parallelization of MAFFT for large-scale
196 multiple sequence alignments. *Bioinformatics*, *34*(14), 2490-2492.
- 197 Pegu, A., O'Connell, S. E., Schmidt, S. D., O'Dell, S., Talana, C. A., Lai, L., . . . Corbett, K. S. (2021).
198 Durability of mRNA-1273 vaccine-induced antibodies against SARS-CoV-2 variants. *Science*,
199 *373*(6561), 1372-1377.
- 200 Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees
201 for large alignments. *PLoS one*, *5*(3), e9490.
- 202 Pulliam, J. R. C., van Schalkwyk, C., Govender, N., von Gottberg, A., Cohen, C., Groome, M. J., . . .
203 Moultrie, H. (2021). Increased risk of SARS-CoV-2 reinfection associated with emergence of
204 the Omicron variant in South Africa. 2021.2011.2011.21266068.
205 doi:10.1101/2021.11.11.21266068 %J medRxiv
- 206 Schmidt, M., Arshad, M., Bernhart, S. H., Hakobyan, S., Arakelyan, A., Loeffler-Wirth, H., & Binder, H.
207 (2021). The evolving faces of the SARS-CoV-2 genome. *Viruses*, *13*(9), 1764.
- 208 Sharma, O., Sultan, A. A., Ding, H., & Triggie, C. R. (2020). A Review of the Progress and Challenges of
209 Developing a Vaccine for COVID-19. *Frontiers in immunology*, *11*, 2413.
- 210 Singh, D., & Soojin, V. Y. (2021). On the origin and evolution of SARS-CoV-2. *Experimental &*
211 *Molecular Medicine*, *53*(4), 537-547.
- 212 Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., . . . Qian, Z. (2020). On the origin and continuing
213 evolution of SARS-CoV-2. *National Science Review*, *7*(6), 1012-1023.

214

215 **Figure and table legends:**

216 **Figure 1: Maximum likelihood whole genome-based phylogeny of SARS-CoV-2 VOCs,**
217 **VOIs and VUMs.** Here, phylogroups (PG-I and PG-II) and clades (alpha, beta, gamma,
218 delta, omicron, mu etc.) are marked with respective colors as indicated. Bootstrap values are
219 represented by the radius of circle at the nodes. Common ancestry of omicron and lambda is
220 marked by black star.

221 **Figure 2: Mutational analysis of omicron.** Six panel image displays the most mutated
222 samples, overall mutations per samples, most frequent events per class of mutation category,
223 changes of nucleotide per type, nucleotide wise most frequent events and protein level most
224 frequent events for the genomes used in the study.

225 **Figure 3: Mutational analysis of omicron (A)** Number of mutations in the coding region is
226 in the centre of the pie-chart representing indels (black), synonymous (yellow) and non-
227 synonymous (red) SNPs. Type and number of mutations in the intergenic region is
228 represented by pie charts blue, light blue and white as represented in the color legends. **(B)**
229 Bar graph representing number of mutations in the genomic region of SARS-CoV-2. **(C)**
230 Some of the top mutations (pl. refer table 3 for all top mutations in omicron) among the
231 omicron variant are represented by stars of black: indels, yellow: synonymous and red: non-
232 synonymous mutations.

233 **Table 1:** Metadata of the VOCs, VOIs and VUMs strains used in the present study.

234 **Table 2:** Genomic region wise mutational count of the omicron isolates by taking
235 NC_045512.2 as a reference.

236 **Table 3:** Top mutations (>185 in count) in omicron variant as compared to the reference
237 sequence NC_045512.2

238 **Supplementary material:**

239 **Supplementary figure 1: Maximum likelihood whole genome-based phylogeny of SARS-**
240 **CoV-2 VOCs, VOIs and VUMs.** Here, phylogroups (PG-I and PG-II) and clades (alpha,
241 beta, gamma, delta, omicron, mu etc.) are marked with respective colors as indicated.
242 Bootstrap values are represented by the radius of circle at the nodes. For better visualisation
243 of tree topology and bootstrap values branch length is ignored.

244 **Supplementary table 1:** Total mutations in strains under study using NC_045512.2 as
245 reference genome.

246 **Supplementary table 2:** Total mutations in strains under study using omicron (OL677199)
247 as reference genome.

Omicron	hCoV-19/G VOC_Omic betacoron: EPI_ISL_69	28-11-2021	Europe	Germany Hesse	Europe	Germany Hesse	genome	29747	Human	38	Male	B.1.1.529	GR	Institute for Infectious Diseases and Hospital Hygiene	2-12-2021
Omicron	hCoV-19/S VOC_Omic betacoron: EPI_ISL_69	24-11-2021	Asia	South Korea	Asia	South Korea	genome	29872	Human	56	Female	B.1.1.529	GR	Division of Infectious Diseases and Hospital Hygiene	2-12-2021
Omicron	hCoV-19/S VOC_Omic betacoron: EPI_ISL_69	25-11-2021	Asia	South Korea	Asia	South Korea	genome	29869	Human	46	Female	B.1.1.529	GR	Division of Infectious Diseases and Hospital Hygiene	2-12-2021
Omicron	hCoV-19/S VOC_Omic betacoron: EPI_ISL_69	29-11-2021	Asia	South Korea	Asia	South Korea	genome	29862	Human	38	Male	B.1.1.529	GR	Division of Infectious Diseases and Hospital Hygiene	2-12-2021
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_67	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	7-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_70	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	14-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_70	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	14-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_70	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Department of COVID-19 (Aminu S. J.)	14-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_70	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Department of COVID-19 (Aminu S. J.)	14-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_70	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	14-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_71	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29861	Human	unknown	unknown	B.1.1.7	GRY	Respiratory COVID-19 (PHE Covid)	18-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_72	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	22-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Department of COVID-19 (Aminu S. J.)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Department of COVID-19 (Aminu S. J.)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Alpha	hCoV-19/E VOC_Alpha betacoron: EPI_ISL_74	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29763	Human	unknown	unknown	B.1.1.7	GRY	Centre for COVID-19 (Angela Bec)	29-12-2020
Beta	hCoV-19/E VOC_Beta betacoron: EPI_ISL_77	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29834	Human	29	Male	B.1.351	GH	Respiratory COVID-19 (PHE Covid)	6-1-2021
Beta	hCoV-19/E VOC_Beta betacoron: EPI_ISL_77	2020	Europe	United Kingdom	Europe	United Kingdom	genome	29803	Human	37	Male	B.1.351	GH	Respiratory COVID-19 (PHE Covid)	6-1-2021
Beta	hCoV-19/A VOC_Beta betacoron: EPI_ISL_85	2020	Europe	Austria	Europe	Austria	genome	29773	Human	unknown	unknown	B.1.351	GH	Austrian Agency for Health and Food Safety	20-1-2021
Beta	hCoV-19/Z VOC_Beta betacoron: EPI_ISL_11	1-9-2020	Africa	Zimbabwe	Africa	Zimbabwe	genome	29779	Human	36	Male	B.1.351	GH	National Microbiology Reference Laboratory	9-3-2021
Beta	hCoV-19/N VOC_Beta betacoron: EPI_ISL_13	2020	Europe	Netherlands	Europe	Netherlands	genome	29757	Human	unknown	unknown	B.1.351	GH	Dutch COVID-19 Reference Laboratory	23-3-2021
Beta	hCoV-19/C VOC_Beta betacoron: EPI_ISL_17	2020-02	Asia	Qatar Doha	Asia	Qatar Doha	genome	29764	Human	62	Male	B.1.351	GH	Ministry of Biomedical Research	25-4-2021
Beta	hCoV-19/C VOC_Beta betacoron: EPI_ISL_22	27-3-2020	Asia	Qatar Doha	Asia	Qatar Doha	genome	29764	Human	28	Male	B.1.351	GH	Ministry of Biomedical Research	26-5-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_23	31-8-2020	Africa	South Africa	Africa	South Africa	genome	29804	Human	21	Male	B.1.351	G	Vaccines at KRISP, KZN	31-5-2021
Beta	hCoV-19/U VOC_Beta betacoron: EPI_ISL_25	7-5-2020	North America	USA South Carolina	North America	USA South Carolina	genome	29800	Human	unknown	Female	B.1.351	GH	Microbiology	7-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	28-8-2020	Africa	South Africa Eastern Cape	Africa	South Africa Eastern Cape	genome	29890	Human	39	Male	B.1.351	GH	National Institute for Communicable Diseases	7-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	2020	Africa	South Africa	Africa	South Africa	genome	29849	Human	unknown	unknown	B.1.351	GH	National Institute for Communicable Diseases	8-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	2020	Africa	South Africa	Africa	South Africa	genome	29852	Human	unknown	unknown	B.1.351	GH	National Institute for Communicable Diseases	8-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	2020	Africa	South Africa	Africa	South Africa	genome	29856	Human	unknown	unknown	B.1.351	GH	National Institute for Communicable Diseases	8-6-2021
Beta	hCoV-19/Z VOC_Beta betacoron: EPI_ISL_24	27-5-2020	Africa	Zimbabwe Masvingo	Africa	Zimbabwe Masvingo	genome	29827	Human	34	Male	B.1.351	GH	National Microbiology Reference Laboratory	11-6-2021
Beta	hCoV-19/Z VOC_Beta betacoron: EPI_ISL_24	9-7-2020	Africa	Zimbabwe Masvingo	Africa	Zimbabwe Masvingo	genome	29847	Human	28	Female	B.1.351	G	National Microbiology Reference Laboratory	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	4-9-2020	Africa	South Africa Gauteng	Africa	South Africa Gauteng	genome	29828	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	8-9-2020	Africa	South Africa Free State	Africa	South Africa Free State	genome	29807	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	9-9-2020	Africa	South Africa Gauteng	Africa	South Africa Gauteng	genome	29831	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	9-9-2020	Africa	South Africa Gauteng	Africa	South Africa Gauteng	genome	29878	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	10-9-2020	Africa	South Africa Gauteng	Africa	South Africa Gauteng	genome	29832	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	10-9-2020	Africa	South Africa Gauteng	Africa	South Africa Gauteng	genome	29832	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/S VOC_Beta betacoron: EPI_ISL_24	10-9-2020	Africa	South Africa Free State	Africa	South Africa Free State	genome	29825	Human	unknown	unknown	B.1.351	GH	Vaccines at KRISP, KZN	11-6-2021
Beta	hCoV-19/R VOC_Beta betacoron: EPI_ISL_47	4-6-2020	Africa	Rwanda Rusizi	Africa	Rwanda Rusizi	genome	29832	Human	41	Male	B.1.351	GH	Rwanda National Institute for Biomedical Research	4-10-2021
Beta	hCoV-19/U VOC_Beta betacoron: EPI_ISL_51	28-8-2020	Asia	United Arab Emirates	Asia	United Arab Emirates	genome	29885	Human	73	Male	B.1.351	GH	BTC, Khalifa Bin Zayed	14-10-2021
Beta	hCoV-19/U VOC_Beta betacoron: EPI_ISL_66	26-6-2020	Africa	South Africa North West	Africa	South Africa North West	genome	29760	Human	78	Female	B.1.351	GH	TSHEPONG National Institute for Communicable Diseases	24-11-2021
Gamma	hCoV-19/F VOC_Gamma betacoron: EPI_ISL_24	23-3-2020	Europe	France Provence-Alpes-Cote d'Azur	Europe	France Provence-Alpes-Cote d'Azur	genome	29673	Human	unknown	unknown	P.1.15	GR	MEPHI, Aix-Marseille University	8-6-2021
Gamma	hCoV-19/U VOC_Gamma betacoron: EPI_ISL_22	7-4-2020	North America	USA Illinois	North America	USA Illinois	genome	29883	Human	unknown	unknown	P.1	GR	Reditus Lal	25-5-2021
Gamma	hCoV-19/U VOC_Gamma betacoron: EPI_ISL_26	23-4-2020	North America	USA Indiana	North America	USA Indiana	genome	29830	Human	93	Male	P.1	GR	ISDH	18-6-2021
Gamma	hCoV-19/U VOC_Gamma betacoron: EPI_ISL_22	22-7-2020	North America	USA Texas Houston	North America	USA Texas	genome	29678	Human	unknown	unknown	P.1.1	GR	Houston Methodist Hospital	21-5-2021
Gamma	hCoV-19/U VOC_Gamma betacoron: EPI_ISL_22	22-7-2020	North America	USA Texas Houston	North America	USA Texas	genome	29771	Human	unknown	unknown	P.1.1	GR	Houston Methodist Hospital	21-5-2021
Gamma	hCoV-19/U VOC_Gamma betacoron: EPI_ISL_22	31-7-2020	North America	USA Texas Houston	North America	USA Texas	genome	29787	Human	unknown	unknown	P.1.1	GR	Houston Methodist Hospital	21-5-2021

Mu	hCoV-19/F VOI_Mu_h betacoron: EPI_ISL_41	2021	Europe	France	Nouvelle-Aquitaine	Europe	France	Nouvelle-A genome	29827	Human	28	Female	B.1.621.1	GH	Cerballanc CERBA Hea Bénédicte https://ww	15-9-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29781	Human	unknown	Male	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29781	Human	unknown	Female	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29781	Human	unknown	Female	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29775	Human	unknown	Female	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29778	Human	unknown	Male	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_56	2021	South Ame	Colombia	Valle del Cauca	South Ame	Colombia	Valle del C; genome	29781	Human	unknown	Female	B.1.621	GH	SYNLAB ÁN Laboratoric Fredy A. Tz https://ww	27-10-2021	
Mu	hCoV-19/U VOI_Mu_h betacoron: EPI_ISL_57	2021	North Ame	USA	Indiana	North Ame	USA	Indiana genome	29799	Human	38	Female	B.1.621	GH	PPHC-Purd Animal Dis Redbecca F https://ww	29-10-2021	
Mu	hCoV-19/C VOI_Mu_h betacoron: EPI_ISL_65	14-10-2020	South Ame	Colombia	Antioquia	Apartado	South Ame	Colombia	Antioquia genome	29781	Human	89	Male	B.1.621	GH	SOMER Laboratoric Idabely Be https://ww	19-11-2021
GH	hCoV-19/Ir VUM_GH_ betacoron: EPI_ISL_31	18-5-2021	Asia	Indonesia	Jambi	Jambi	Asia	Indonesia	Jambi genome	29782	Human	18	Female	B.1.640	GH	RSUD Rade Eijkman In: Sukma Okt https://ww	1-8-2021
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_55	28-9-2021	Africa	Republic of	Brazzaville	Africa	Republic o	Brazzaville genome	29683	Human	39	Female	B.1.640	GH	Fondation Fondation Mfoutou N https://ww	27-10-2021	
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_64	2-10-2021	North Ame	Canada	Quebec	North Ame	Canada	Quebec genome	29734	Human	34	unknown	B.1.640	GH	Laboratoirc Laboratoirc Sandrine N https://ww	18-11-2021	
GH	hCoV-19/S VUM_GH_ betacoron: EPI_ISL_55	6-10-2021	Europe	Switzerland	Vaud	Europe	Switzerland	Vaud genome	29788	Human	unknown	unknown	B.1.640	GH	Viollier AG Departmer CHRISTIAN B https://ww	26-10-2021	
GH	hCoV-19/E VUM_GH_ betacoron: EPI_ISL_54	11-10-2021	Europe	United Kin	England	Europe	United Kin	England genome	29779	Human	unknown	unknown	B.1.640	GH	Respirator COVID-19 (PHE Covid https://ww	25-10-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	12-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29799	Human	9	Female	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	12-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29796	Human	9	Male	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	12-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29821	Human	10	Male	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_59	12-10-2021	Europe	France	Provence-/	Alpes-Mari	Europe	France	Provence-/ genome	29728	Human	36	Male	B.1.640	GH	LABORATC CNR Virus (Antonin Bz https://ww	4-11-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	14-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29770	Human	10	Male	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/E VUM_GH_ betacoron: EPI_ISL_55	16-10-2021	Europe	France	Ile-de-France	Europe	France	Ile-de-Fran genome	29822	Human	36	Female	B.1.640	G	LBM Porte CERBA Hea Bénédicte https://ww	27-10-2021	
GH	hCoV-19/S VUM_GH_ betacoron: EPI_ISL_56	18-10-2021	Europe	United Kin	Scotland	Europe	United Kin	Scotland genome	29434	Human	unknown	unknown	B.1.640	GH	Lighthouse Wellcome : Harper Var https://ww	27-10-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	18-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29770	Human	10	Male	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	18-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29829	Human	10	Male	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	18-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29733	Human	5	Female	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_60	19-10-2021	Africa	Republic of	Brazzaville	Africa	Republic o	Brazzaville genome	29714	Human	32	Female	B.1.640	GH	Fondation Fondation Mfoutou N https://ww	9-11-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_56	20-10-2021	Europe	France	Bretagne	Bannalec	Europe	France	Bretagne genome	29770	Human	12	Female	B.1.640	GH	Laboratoirc CHU Pontc GROLHIER https://ww	28-10-2021
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_60	20-10-2021	Africa	Republic of	Brazzaville	Africa	Republic o	Brazzaville genome	29649	Human	4	Female	B.1.640	GH	Fondation Fondation Mfoutou N https://ww	9-11-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_64	21-10-2021	Europe	France	Bretagne	Vannes	Europe	France	Bretagne genome	29733	Human	16	Female	B.1.640	GH	Labo Analy National Rr Marion Bai https://ww	18-11-2021
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_62	22-10-2021	Europe	France	Ile-de-France	Europe	France	Ile-de-Fran genome	29741	Human	76	Male	B.1.640	GH	Hôpital Loj Departmer Christophe https://ww	12-11-2021	
GH	hCoV-19/It VUM_GH_ betacoron: EPI_ISL_57	23-10-2021	Europe	Italy	Lazio	Rome	Europe	Italy	Lazio genome	29774	Human	unknown	Male	B.1.640	GH	IRCCS San r IRCCS Regi Frauke Go https://ww	31-10-2021
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_60	23-10-2021	Africa	Republic of	Brazzaville	Africa	Republic o	Brazzaville genome	29714	Human	43	Female	B.1.640	GH	Fondation Fondation Mfoutou N https://ww	9-11-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_59	25-10-2021	Europe	France	Ile-de-France	Europe	France	Ile-de-Fran genome	29814	Human	61	Male	B.1.640	GH	LBM Bois d CERBA Hea Bénédicte https://ww	4-11-2021	
GH	hCoV-19/C VUM_GH_ betacoron: EPI_ISL_60	25-10-2021	Africa	Republic of	Brazzaville	Africa	Republic o	Brazzaville genome	29683	Human	35	Male	B.1.640	GH	Fondation Fondation Mfoutou N https://ww	9-11-2021	
GH	hCoV-19/F VUM_GH_ betacoron: EPI_ISL_63	25-10-2021	Europe	France	Ile-de-Fran	Lisses	Europe	France	Ile-de-Fran genome	29745	Human	61	Male	B.1.640	GH	Labo Analy National Rr Marion Bai https://ww	16-11-2021
Lambda	hCoV-19/E VOI_Lambi betacoron: EPI_ISL_16	2021	South Ame	Ecuador	Guayas	South Ame	Ecuador	Guayas genome	29783	Human	unknown	unknown	C.37	GR	Omics Scie Omics Scie Derly Andr https://ww	19-4-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_59	21-7-2020	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29769	Human	38	Male	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	5-11-2021	
Lambda	hCoV-19/A VOI_Lambi betacoron: EPI_ISL_21	8-11-2020	South Ame	Argentina	Ciudad Autonoma de	South Ame	Argentina	Ciudad Aut genome	29792	Human	27	Female	C.37	GR	Servicio Vir Instituto N Baumeiste https://ww	19-5-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_15	22-12-2020	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29496	Human	34	Female	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	11-4-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_16	1-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29744	Human	unknown	unknown	C.37	GR	Instituto di Laboratoric Lenin Mati https://ww	18-4-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_34	8-1-2021	South Ame	Peru	San Martin	South Ame	Peru	San Martin genome	29791	Human	22	Male	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	14-8-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_34	9-1-2021	South Ame	Peru	San Martin	South Ame	Peru	San Martin genome	29786	Human	25	Male	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	14-8-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_34	11-1-2021	South Ame	Peru	San Martin	South Ame	Peru	San Martin genome	29783	Human	40	Female	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	14-8-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	12-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29496	Human	47	Female	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	4-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	12-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29496	Human	38	Female	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	4-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	12-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29496	Human	74	Female	C.37	GR	Laboratoric Laboratoric Carlos Pad https://ww	4-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Callao	South Ame	Peru	Callao genome	29901	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29890	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29901	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29898	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29895	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29899	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29863	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29895	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29901	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29887	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29901	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29874	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	17-1-2021	South Ame	Peru	Lima	South Ame	Peru	Lima genome	29901	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	
Lambda	hCoV-19/P VOI_Lambi betacoron: EPI_ISL_11	18-1-2021	South Ame	Peru	Callao	South Ame	Peru	Callao genome	29888	Human	unknown	unknown	C.37	GR	Laboratoric Laboratoric Ronnie Ga https://ww	1-3-2021	

Table 2: Genomic region wise mutational count of the omicron isolates by taking reference sequence NC_045512.2

Genomic region	Mutational count	Annotation
5'UTR	309	5' Untranslated region
NSP1	5	
NSP2	31	
NSP3	1572	
NSP4	325	
NSP5	317	
NSP6	595	
NSP7	0	
NSP8	2	
NSP9	9	
NSP10	301	
NSP11	0	
NSP12a	0	
NSP12b	632	
NSP13	14	
NSP14	319	
NSP15	6	
NSP16	14	RNA dependent RNA polymerase
S	10658	Spike
ORF3a	313	ORF3a protein
E	296	Envelope
M	850	Membrane
ORF6	303	ORF6 protein
ORF7a	2	ORF7a protein
ORF7b	311	ORF7b protein
ORF8	4	ORF8 protein
N	823	Nucleocapsid protein
ORF10	1	ORF10 protein
3'UTR	249	3' Untranslated region

Table 3: Top mutations (>185 count) in omicron variant as compared to the reference sequence NC_045512.2

annotator	protein	variant	varclass	Count	refpos	refvar	qvar	qpos	qlength
Spike	S	A67	deletion_fi	575	21762	C	.	21483	29387
Predicted Ꞥ	NSP3	A1892T	SNP	302	8393	G	A	8124	29387
Transmem	NSP6	I189V	SNP	302	11537	A	G	11259	29387
RNA-deper	NSP12b	P314L	SNP	302	14408	C	T	14130	29387
Spike	S	T547K	SNP	302	23202	C	A	22915	29387
Spike	S	D614G	SNP	302	23403	A	G	23116	29387
Spike	S	H655Y	SNP	302	23525	C	T	23238	29387
ORF6 prot	ORF6	M19M	SNP_silent	302	27259	A	C	26972	29387
Predicted Ꞥ	NSP3	K38R	SNP	301	2832	A	G	2566	29387
Spike	S	N679K	SNP	301	23599	T	G	23312	29387
Transmem	NSP4	T492I	SNP	301	10029	C	T	9760	29378
Nucleocap	N	RG203K*	SNP	301	28881	GGG	AAT	28806	29693
Growth-fac	NSP10	V57V	SNP_silent	300	13195	T	C	12917	29387
Spike	S	P681H	SNP	300	23604	C	A	23317	29387
Spike	S	D796Y	SNP	300	23948	G	T	23661	29387
Spike	S	N856K	SNP	300	24130	C	A	23843	29387
Spike	S	Q954H	SNP	300	24424	A	T	24137	29387
Nucleocap	N	RG203KR	SNP	300	28881	GGG	AAC	28594	29387
RNA-deper	NSP12b	N591N	SNP_silent	298	15240	C	T	14962	29387
Spike	S	T95I	SNP	298	21846	C	T	21562	29387
Predicted Ꞥ	NSP3	F106F	SNP_silent	297	3037	C	T	2771	29387
Spike	S	G339D	SNP	297	22578	G	A	22291	29387
ORF3a pro	ORF3a	T64T	SNP_silent	297	25584	C	T	25297	29387
NA	5'UTR	241	extragenic	297	241	C	T	187	29693
3C-like pro	NSP5	P132H	SNP	296	10449	C	A	10180	29387
3'-to-5' exc	NSP14	I42V	SNP	296	18163	A	G	17885	29387
Envelope	E	T9I	SNP	296	26270	C	T	25983	29387
ORF7b pro	ORF7b	L17L	SNP_silent	296	27807	C	T	27520	29387
Spike	S	N969K	SNP	294	24469	T	A	24182	29387
Predicted Ꞥ	NSP3	A889A	SNP_silent	293	5386	T	G	5120	29387
Spike	S	L981F	SNP	292	24503	C	T	24216	29387
Spike	S	D1146D	SNP_silent	292	25000	C	T	24713	29387
Membran	M	A63T	SNP	289	26709	G	A	26422	29387
Predicted Ꞥ	NSP3	S1265	deletion	288	6513	GTT	.	6246	29387
Transmem	NSP6	L105	deletion	287	11286	TGTCTGGT	.	11016	29387

Spike	S	I68	deletion_fi	287	21767	CATG	.	21486	29387
Spike	S	E484A	SNP	284	23013	A	C	22726	29387
Spike	S	S477N	SNP	283	22992	G	A	22705	29387
Spike	S	T478K	SNP	283	22995	C	A	22708	29387
Spike	S	Q493R	SNP	282	23040	A	G	22753	29387
Spike	S	Q498R	SNP	281	23055	A	G	22768	29387
Spike	S	N501Y	SNP	281	23063	A	T	22776	29387
Spike	S	G496S	SNP	280	23048	G	A	22761	29387
Spike	S	Y505H	SNP	277	23075	T	C	22788	29387
Membrane	M	D3G	SNP	275	26530	A	G	26243	29387
Membrane	M	Q19E	SNP	272	26577	C	G	26290	29387
Spike	S	S371L	SNP	270	22673	TC	CT	22386	29387
Spike	S	S373P	SNP	270	22679	T	C	22392	29387
Spike	S	G142	deletion	260	21987	GTGTTTAT	.	21702	29387
Spike	S	S375F	SNP	260	22686	C	T	22399	29387
ORF7b pro	ORF7b	E3*	SNP_stop	253	27762	G	T	27687	29752
Spike	S	I210	insertion_f	243	22193	.	T	21901	29387
Spike	S	R214	insertion_f	243	22203	.	A	21916	29387
Spike	S	R214R	SNP_silent	243	22204	T	A	21917	29387
Nucleocap	N	E31	deletion	243	28362	GAGAACG	(.	28074	29378
Spike	S	L212*	SNP_stop	243	22197	T	G	22118	29749
Spike	S	N211K	SNP	242	22195	T	G	21903	29387
Spike	S	L212C	SNP	242	22197	TA	GC	21905	29387
Spike	S	S214	insertion	242	22201	.	AGC	21910	29387
Spike	S	V213	insertion_f	242	22202	.	A	21914	29387
NA	3'UTR	28271	extragenic	242	28271	A	T	27984	29378
Nucleocap	N	P13L	SNP	241	28311	C	T	28024	29378
Spike	S	N764K	SNP	234	23854	C	A	23567	29387
Spike	S	G446S	SNP	203	22898	G	A	22611	29387
Spike	S	N440K	SNP	199	22882	T	G	22595	29387
Spike	S	K417N	SNP	183	22813	G	T	22526	29387

Variants of SARS-CoV-2

- VOC Gamma
- VOC Beta
- VOI Mu
- VUM GH
- VOC Delta
- VOC Alpha
- VOI Lambda
- VOC Omicron

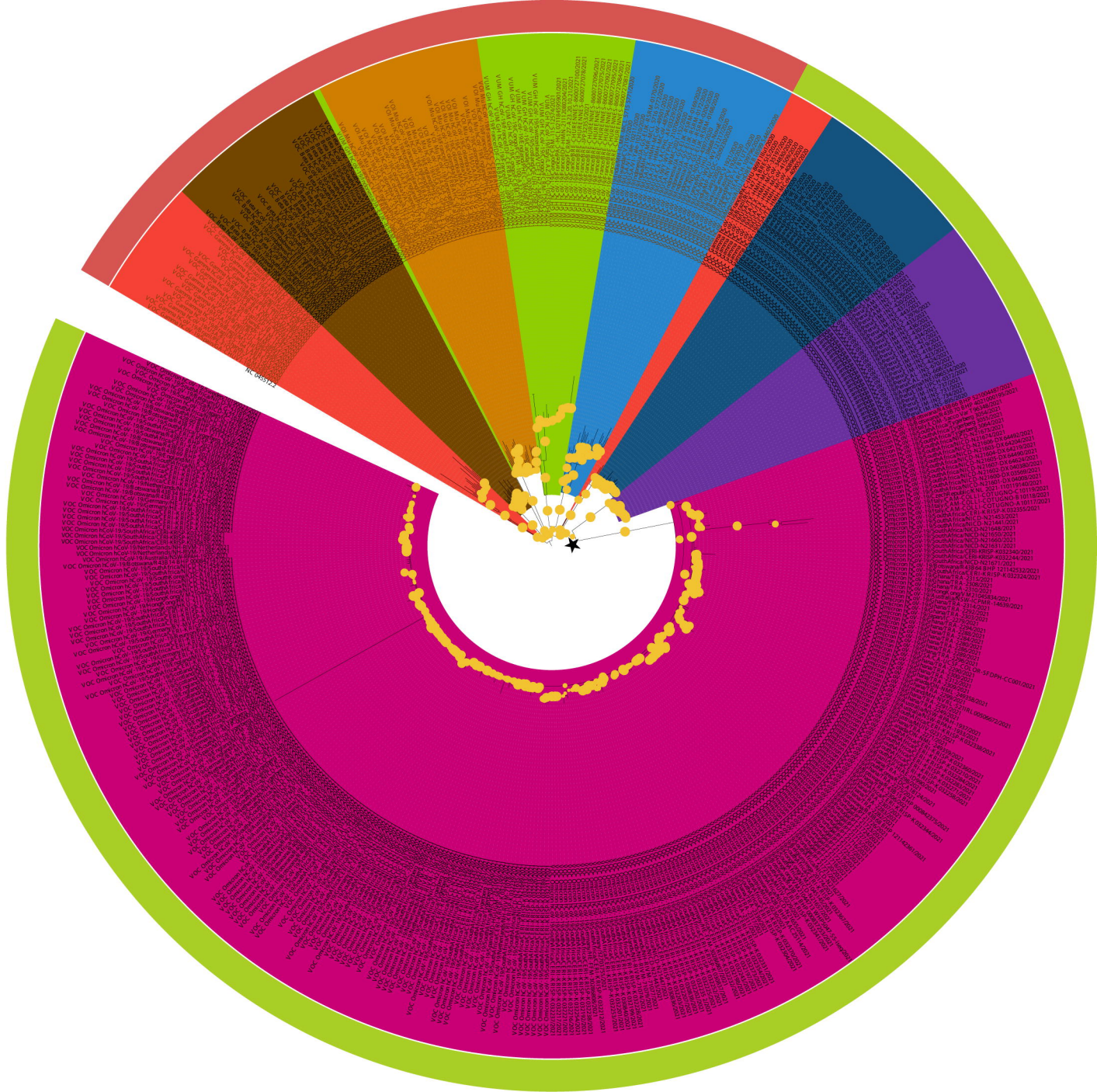
Phylogroups

- PG-I
- PG-II

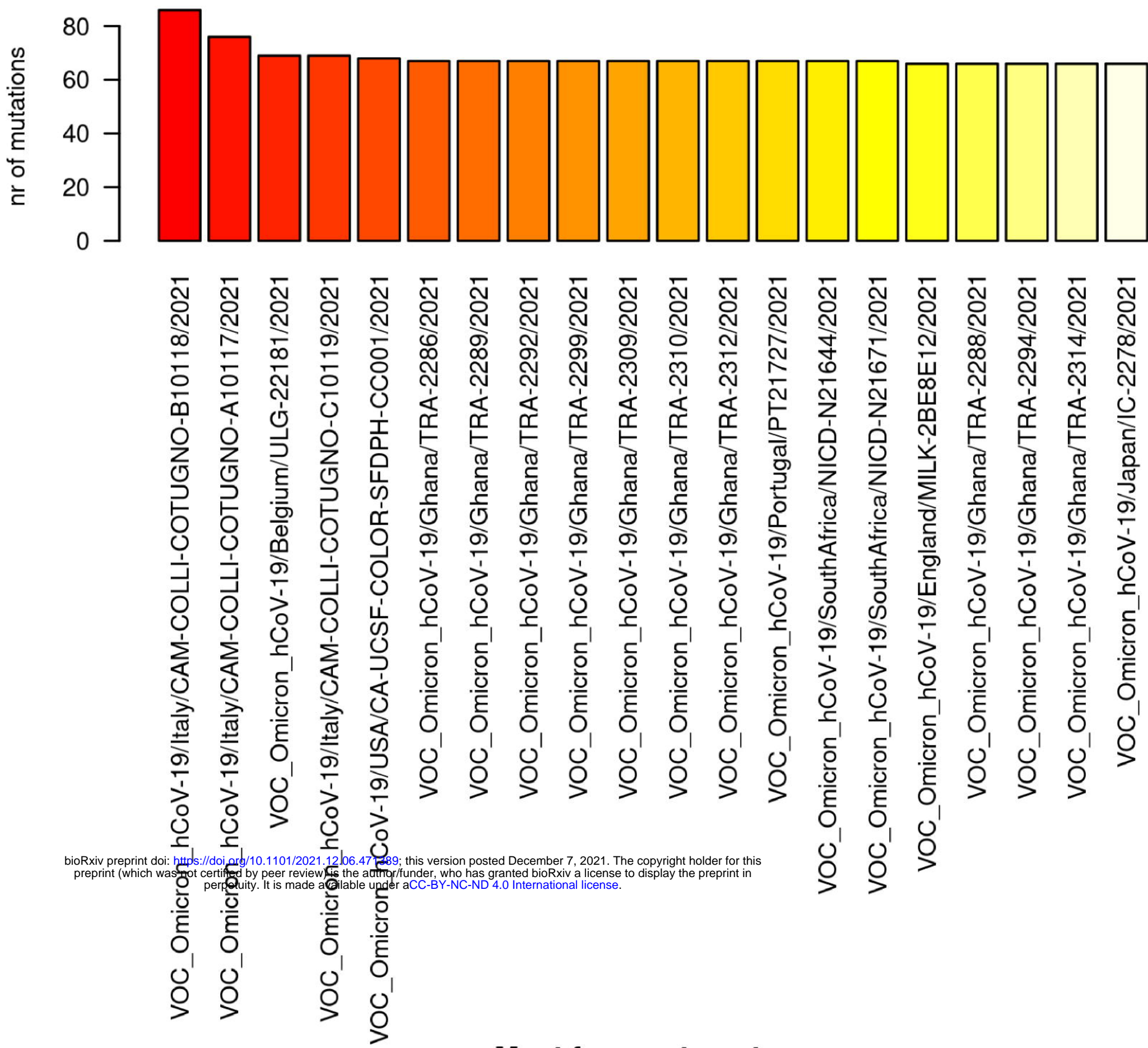
bootstrap

- 0
- 0.25
- 0.5
- 0.75
- 1

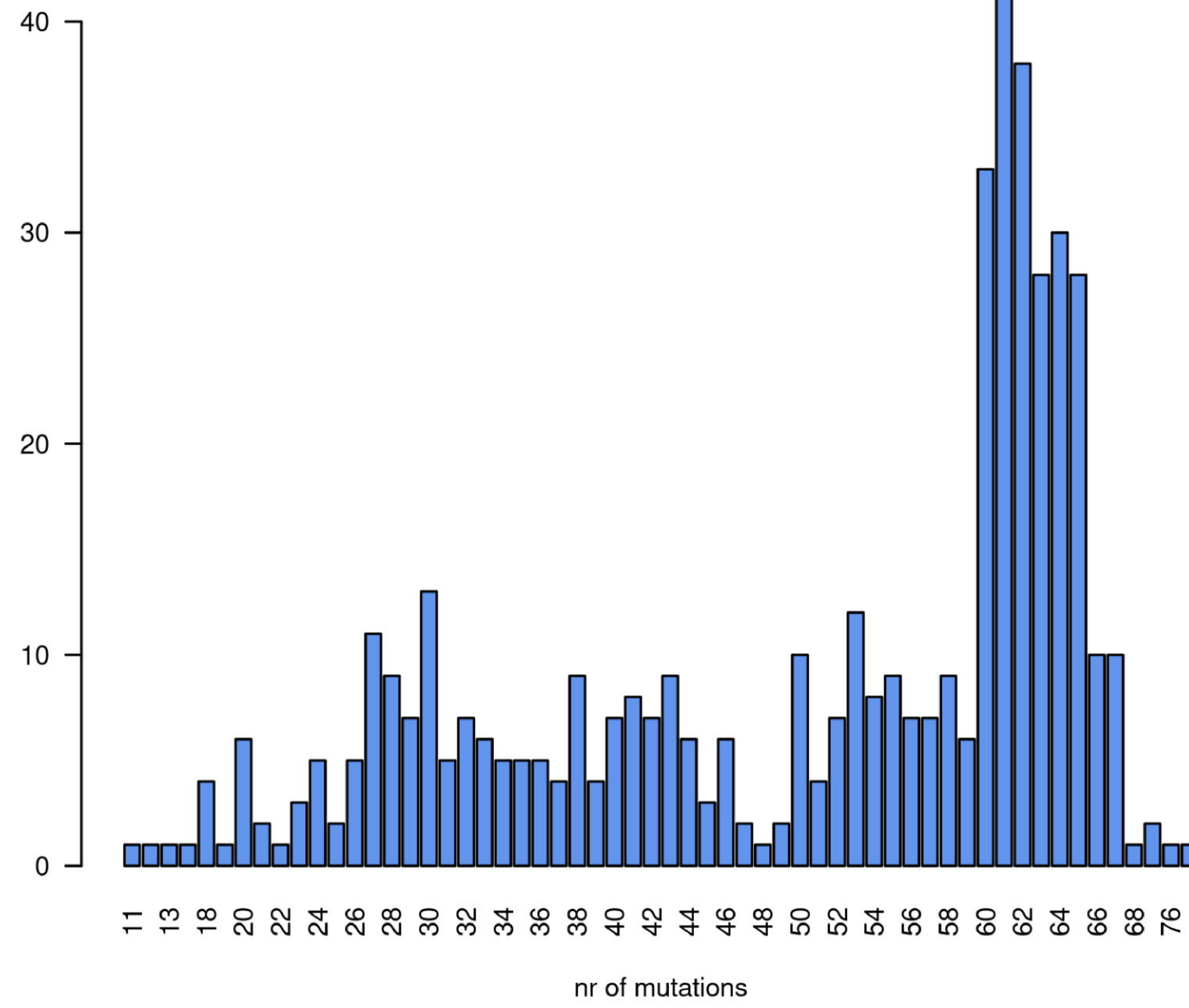
Tree scale: 0.001



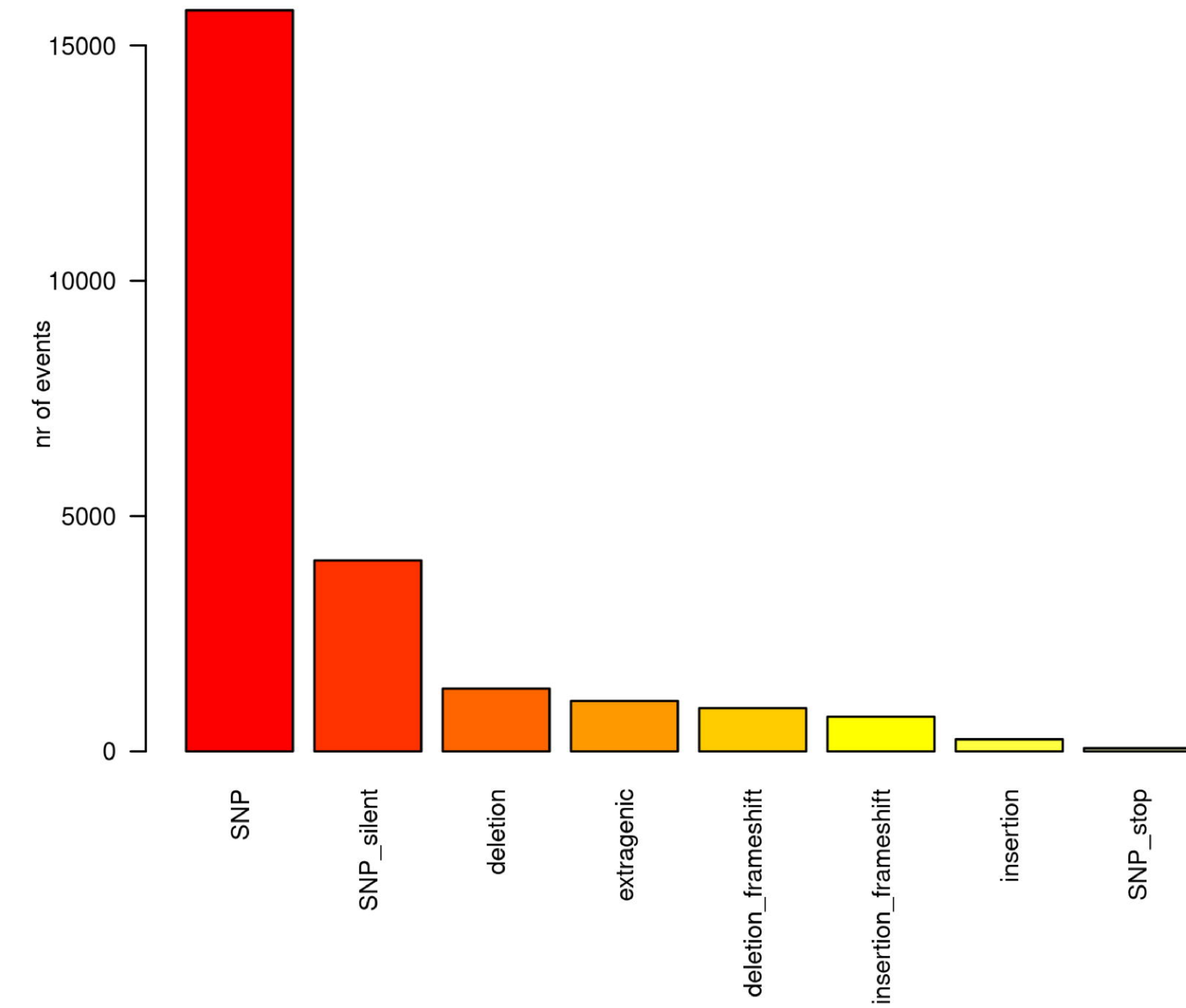
Most mutated samples



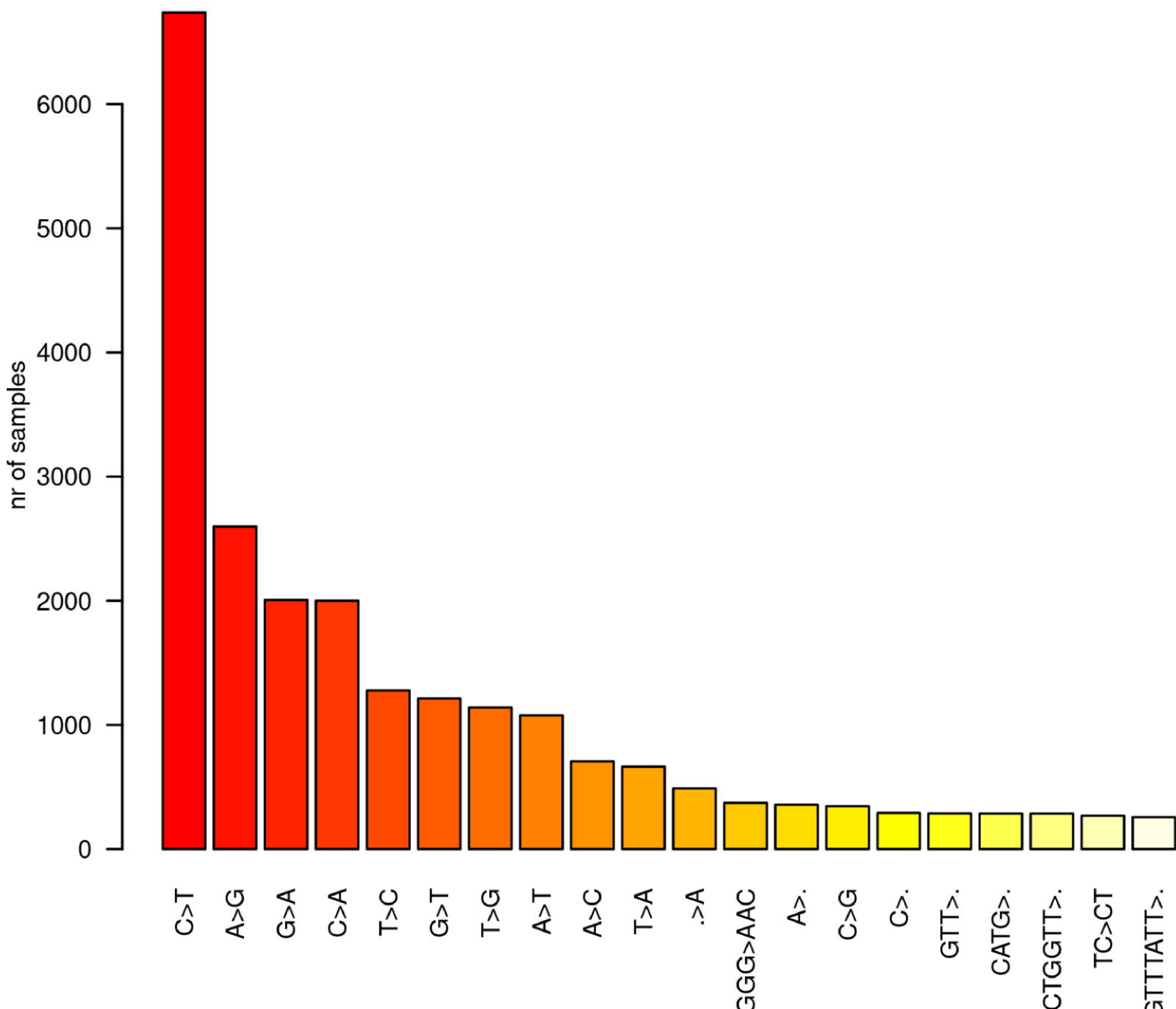
Overall mutations per sample



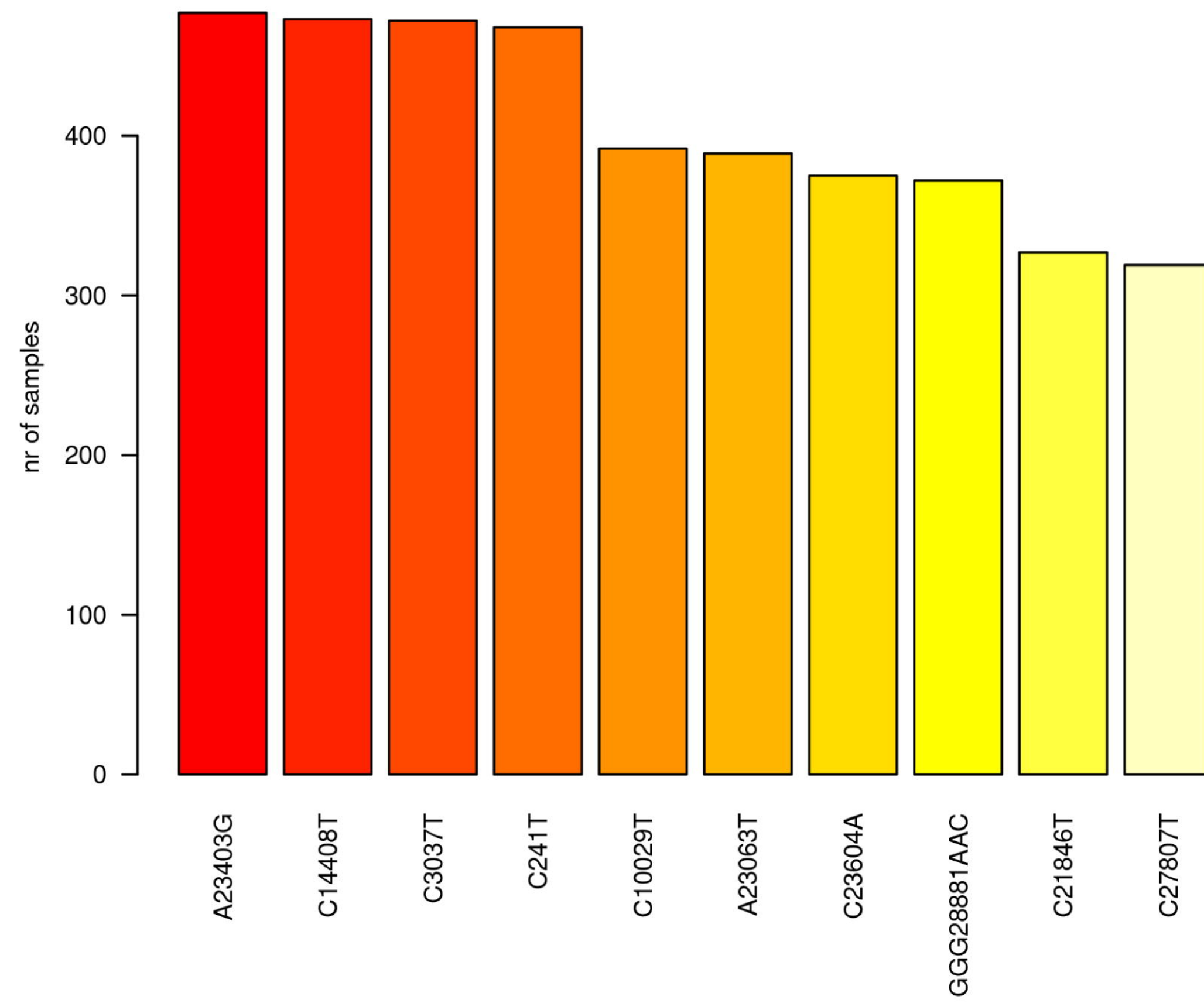
Most frequent events per class



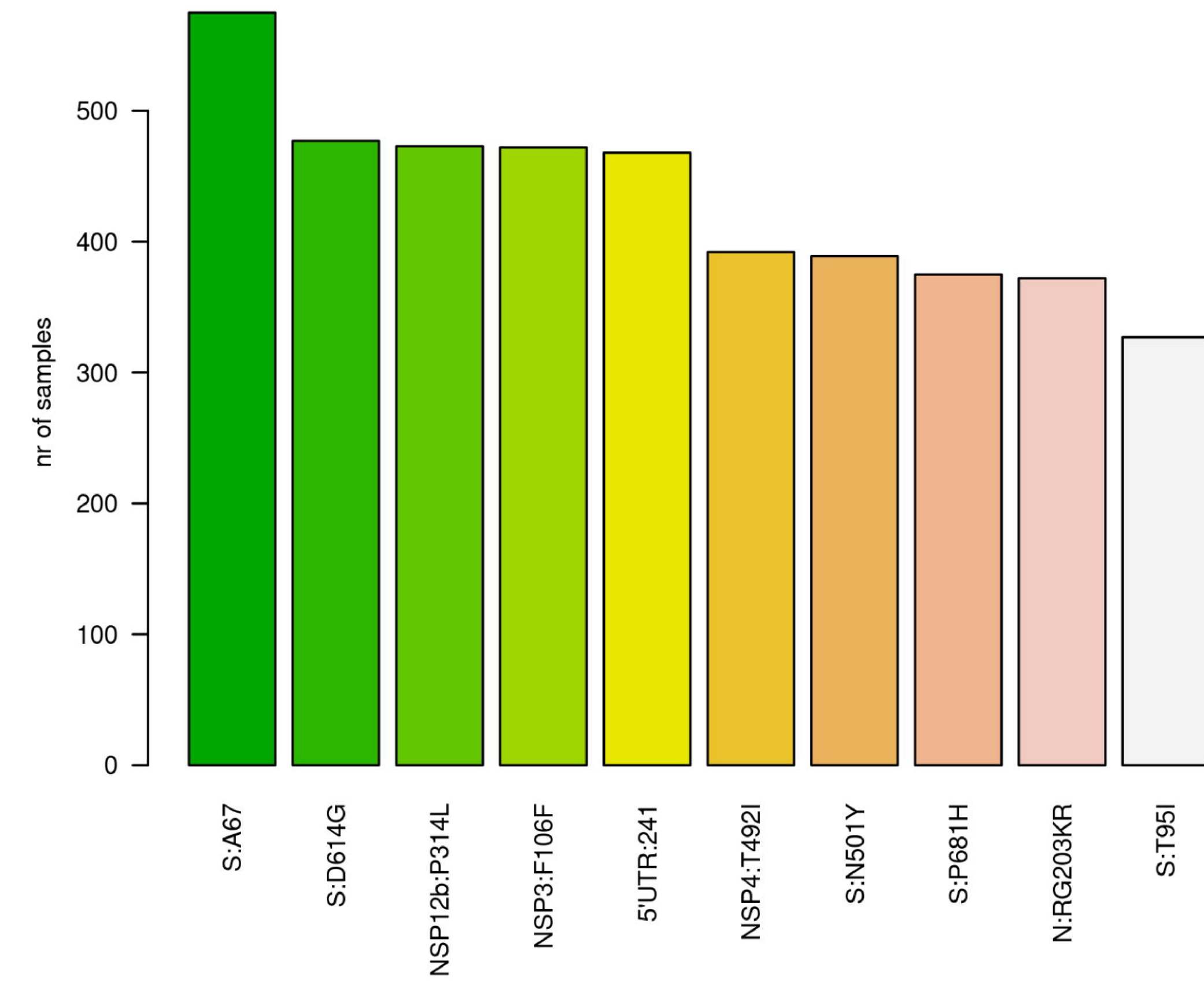
Most frequent per type



Most frequent events (nucleotide)



Most frequent events (protein)



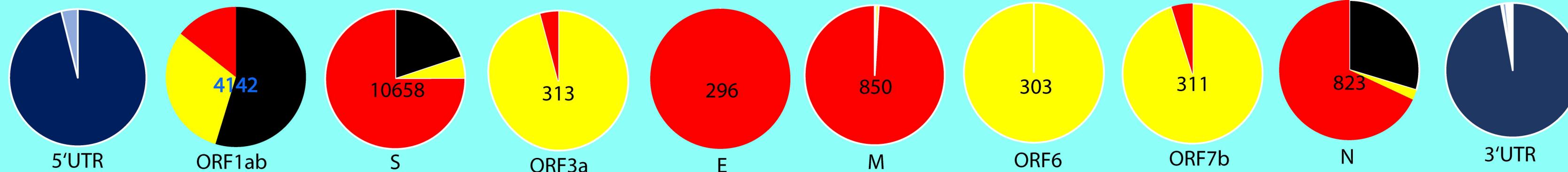
■ 5'UTR 241
 □ 5'UTR 21

■ Insertion/Deletion

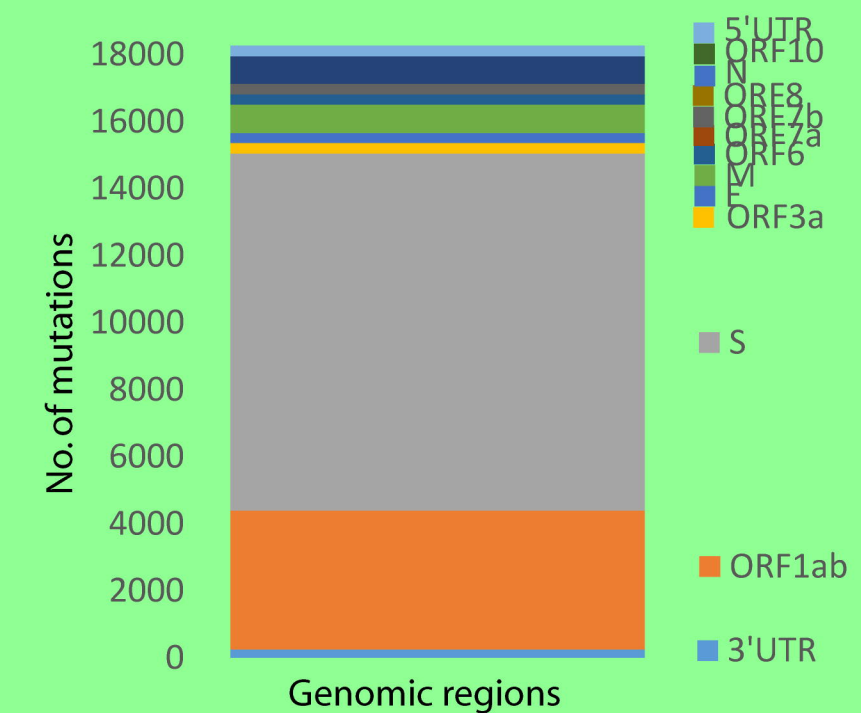
■ Synonymous SNP

■ Non-synonymous SNP

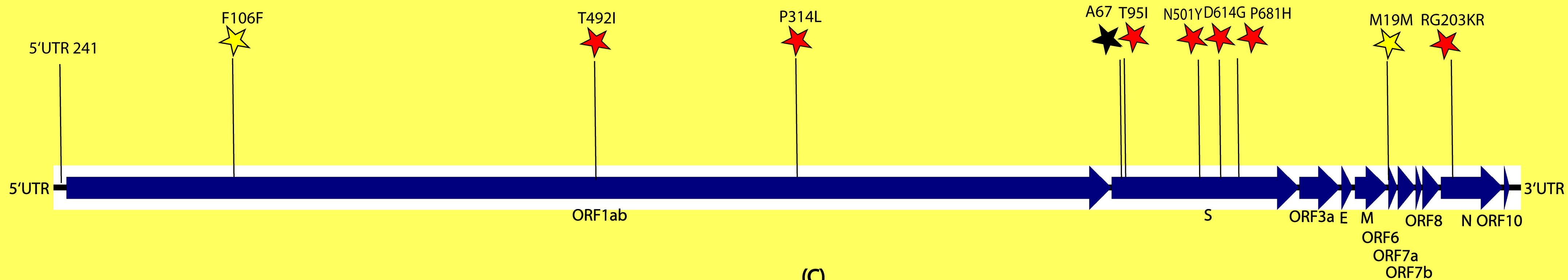
■ 3'UTR 28271
 29705, 28273, 29844,
 29750, 29867, 29742



(A)



(B)



(C)