

## **A potential SARS-CoV-2 variant of interest (VOI) harboring mutation E484K in the Spike protein was identified within lineage B.1.1.33 circulating in Brazil**

Paola Cristina Resende<sup>1\*</sup>; Tiago Gräf<sup>2\*</sup>; Anna Carolina Dias Paixão<sup>1</sup>; Luciana Appolinario<sup>1</sup>; Renata Serrano Lopes<sup>1</sup>; Ana Carolina da Fonseca Mendonça<sup>1</sup>; Alice Sampaio Barreto da Rocha<sup>1</sup>; Fernando Couto Motta<sup>1</sup>; Lidio Gonçalves Lima Neto<sup>3</sup>; Ricardo Khouri<sup>2,4</sup>; Camila Indiani de Oliveira<sup>2,4</sup>; Pedro Santos-Muccillo<sup>2,5</sup>; Joao Felipe Bezerra<sup>6</sup>; Dalane Loudal Florentino Teixeira<sup>7</sup>; Irina Riediger<sup>8</sup>; Maria do Carmo Debur<sup>8</sup>; Rodrigo Ribeiro-Rodrigues<sup>9</sup>; Anderson Brandao Leite<sup>10</sup>; Cliomar Alves do Santos<sup>11</sup>; Tatiana Schäffer Gregianini<sup>12</sup>; Sandra Bianchini Fernandes<sup>13</sup>; André Felipe Leal Bernardes<sup>14</sup>; Andrea Cony Cavalcanti<sup>15</sup>; Fábio Miyajima<sup>16</sup>; Claudio Sachhi<sup>17</sup>; Tirza Mattos<sup>18</sup>; Cristiano Fernandes da Costa<sup>19</sup>; Edson Delatorre<sup>20\*\*</sup>; Gabriel L Wallau<sup>21\*\*</sup>; Felipe G Naveca<sup>22\*\*</sup>; Gonzalo Bello<sup>23\*\*</sup>, Marilda Mendonça Siqueira<sup>1\*\*</sup> on behalf of Fiocruz COVID-19 Genomic Surveillance Network

\* These authors share the first authorship

\*\* These authors share the senior authorship

1. Laboratory of Respiratory Viruses and Measles (LVRS), Oswaldo Cruz Institute, Fiocruz
2. Plataforma de Vigilância Molecular, Instituto Gonçalo Moniz, Fiocruz-BA
3. Laboratório Central de Saúde Pública do Estado do Maranhão (LACEN-MA)
4. Laboratório de Enfermidades Infecciosas Transmitidas por Vetores, Instituto Gonçalo Moniz, Fiocruz Bahia
5. Laboratório de Patologia e Biologia Molecular, Instituto Gonçalo Moniz, Fiocruz Bahia
6. Universidade Federal da Paraíba (UFPB)
7. Laboratório Central de Saúde Pública do Estado da Paraíba (LACEN-PB)
8. Laboratório Central de Saúde Pública do Estado do Paraná (LACEN-PR)
9. Laboratório Central de Saúde Pública do Estado do Espírito Santo (LACEN-ES)
10. Laboratório Central de Saúde Pública do Estado do Alagoas (LACEN-AL)
11. Laboratório Central de Saúde Pública do Estado do Sergipe (LACEN-SE)
12. Laboratório Central de Saúde Pública do Estado do Rio Grande do Sul (LACEN-RS)
13. Laboratório Central de Saúde Pública do Estado de Santa Catarina (LACEN-SC)
14. Laboratório Central de Saúde Pública do Estado de Minas Gerais (LACEN-MG)
15. Laboratório Central de Saúde Pública do Estado do Rio de Janeiro (LACEN-RJ)
16. Fundação Oswaldo Cruz - Fiocruz Ceará
17. Instituto Adolfo Lutz (IAL)
18. Laboratório Central de Saúde Pública do Amazonas, Manaus, Amazonas, Brazil.
19. Fundação de Vigilância em Saúde do Amazonas, Manaus, Amazonas, Brazil.
20. Departamento de Biologia. Centro de Ciências Exatas, Naturais e da Saúde, Universidade Federal do Espírito Santo, Alegre, Brazil.
21. Instituto Ageu Magalhaes, Fiocruz-PE
22. Laboratório de Ecologia de Doenças Transmissíveis na Amazônia (EDTA), Leônidas e Maria Deane Institute, Fiocruz
23. Laboratório de AIDS e Imunologia Molecular, Oswaldo Cruz Institute, Fiocruz.

## Summary

The SARS-CoV-2 epidemic in Brazil was dominated by two lineages designated as B.1.1.28 and B.1.1.33. Two SARS-CoV-2 variants harboring mutations at the receptor-binding domain of the Spike (S) protein, designated as lineages P.1 and P.2, evolved within lineage B.1.1.28 and are rapidly spreading in Brazil. Lineage P.1 is considered a Variant of Concern (VOC) because of the presence of multiple mutations in the S protein (including K417T, E484K, N501Y), while lineage P.2 only harbors mutation S:E484K and is considered a Variant of Interest (VOI). Here we report the identification of a new SARS-CoV-2 VOI within lineage B.1.1.33 that also harbors mutation S:E484K and was detected in Brazil between November 2020 and February 2021. This VOI displayed four non-synonymous lineage-defining mutations (NSP3:A1711V, NSP6:F36L, S:E484K, and NS7b:E33A) and was designated as lineage N.9. The VOI N.9 probably emerged in August 2020 and has spread across different Brazilian states from the Southeast, South, North and Northeast regions.

## Background

The SARS-CoV-2 epidemic in Brazil was mainly driven by lineages B.1.1.28 and B.1.1.33 that probably emerged in February 2020 and were the most prevalent variants in most country regions until October 2020 <sup>1,2</sup>. Recent genomic studies, however, bring attention to the emergence of new SARS-CoV-2 variants in Brazil harboring mutations at the receptor-binding site (RBD) of the Spike (S) protein that might impact viral fitness and transmissibility.

So far, one variant of concern (VOC), designated as lineage P.1, and one variant of interest (VOI), designated as lineage P.2, have been identified in Brazil and both evolved from the lineage B.1.1.28. The VOC P.1, first described in January 2021 <sup>3</sup>, displayed an unusual number of lineage-defining mutations in the S protein (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y, T1027I) and its emergence was associated with a second COVID-19 epidemic wave in the Amazonas state <sup>4,5</sup>. The VOI P.2, first described in samples from October 2020 in the state of Rio de Janeiro, was distinguished by the presence of the S:E484K mutation in RBD and other four lineage-defining mutations outside the S protein <sup>6</sup>. The P.2 lineage has been detected as the most

prevalent variant in several states across the country in late 2020 and early 2021 (<https://www.genomahcov.fiocruz.br>).

Several B.1.1.33-derived lineages are currently defined by the Pangolin system including: lineage N.1 detected in the US, lineage N.2 detected in Suriname and France, lineage N.3 circulating in Argentina, and lineages N.4 and B.1.1.314 circulating in Chile. However, none of these B.1.1.33-derived lineages were characterized by mutations of concern in the S protein. Here, we define the lineage N.9 within B.1.1.33 diversity that harbors mutation E484K in the S protein as was detected in different Brazilian states between November, 2020 and February, 2021.

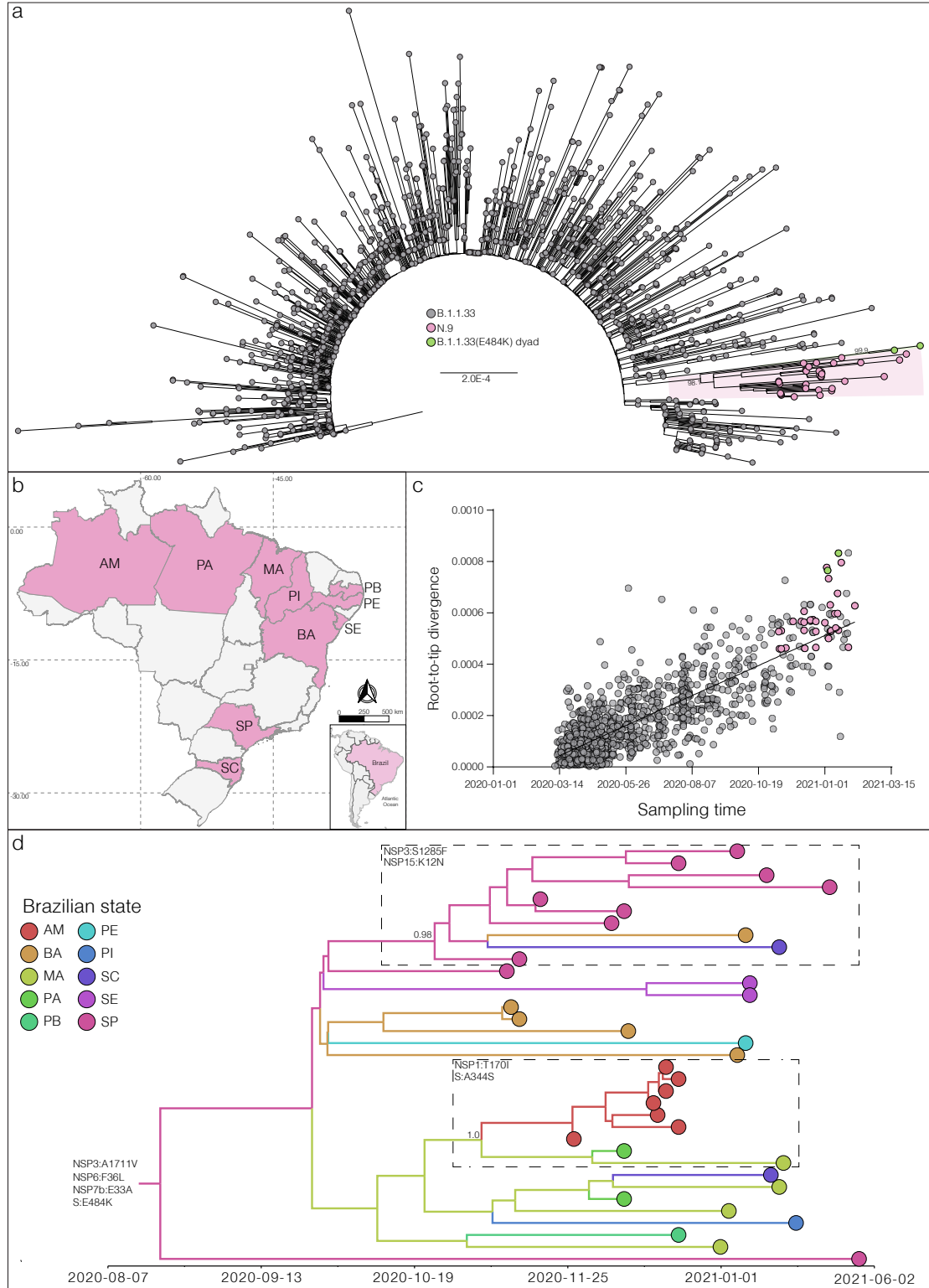
### The Study

Our genomic survey of SARS-CoV-2 positive samples sequenced by the Fiocruz COVID-19 Genomic Surveillance Network between 12th March 2020 and 27th January 2021, identified 422 sequences as belonging to the lineage B.1.1.33 (Appendix 1). Next, these sequences were combined with 816 B.1.1.33 Brazilian genomes available in the EpiCoV database in GISAID by 1<sup>st</sup> March 2021 (Appendix Table 1). Mutational profile was investigated using the nextclade tool (<https://clades.nextstrain.org>), finding the S:E484K mutation in 34 sequences. Maximum-likelihood (ML) phylogenetic analysis conducted using IQ-TREE v2.1.2<sup>7</sup> revealed that 32 B.1.1.33(E484K) sequences branched in a highly supported (approximate likelihood-ratio test [aLRT] = 98%) monophyletic clade that define a potential new VOI designated as N.9 PANGO lineage<sup>8</sup>. The other two sequences harbouring the E484K mutation branched separately in a highly supported (aLRT = 100%) dyad (**Figure 1a**). The VOI N.9 is characterized by four non-synonymous lineage-defining mutations (NSP3:A1711V, NSP6:F36L, S:E484K and NSP7b:E33A) and also contains a group of three B.1.1.33 sequences from the Amazonas state that has no sequencing coverage in the position 484 of the S protein, but share the remaining N.9 lineage-defining mutations (**Table 1**). The B.1.1.33(E484K) dyad comprises two sequences from the Maranhao state and were characterized by a different set of non-synonymous mutations (**Data not shown**).

**Table 1.** Synapomorphic mutations of SARS-CoV-2 lineage N.9.

Genomic region (protein)	Nucleotide	Amino acid
ORF1a	G1264T	-
ORF1a	C7600T	-
ORF1a (NSP3)	C7851T	A2529V (A1711V)
ORF1a (NSP6)	T11078C	F3605L (F36L)
Spike (S)	G23012A	E484K
ORF7b (NSP7b)	A27853C	E33A

Among the 35 genomes identified so far as VOI N.9, 10 Brazilian states were represented, suggesting that this lineage is already highly dispersed in the country. The VOI N.9 was first detected in Sao Paulo state on November 11th, 2020, and soon later in other Brazilian states from the South (Santa Catarina), North (Amazonas and Para) and Northeast (Bahia, Maranhao, Paraiba, Pernambuco, Piaui, and Sergipe) regions (**Figure 1b**). Analysis of the temporal structure using TempEst<sup>9</sup> revealed that the overall divergence of lineage N.9 is consistent with the substitution pattern of other B.1.1.33 sequences (**Figure 1c**), thus suggesting no unusual accumulation of mutations in this VOI. Bayesian reconstruction using a strict molecular clock model with a uniform substitution rate prior ( $8 - 10 \times 10^{-4}$  substitutions/site/year) as implemented in BEAST 1.10<sup>10</sup> estimated the emergence of the VOI N.9 most probably in the states of Sao Paulo (*Posterior State Probability* [*PSP*] = 0.42), Bahia (*PSP* = 0.32) or Maranhao (*PSP* = 0.18) at 15<sup>th</sup> August, 2020 (95% High Posterior Density [HPD]: 16<sup>th</sup> June – 22<sup>th</sup> September, 2020) (**Figure 1d**). This analysis also revealed that some additional mutations were acquired during evolution of VOI N.9 in Brazil, determining two highly supported ( $PP > 0.95$ ) subclades. One subclade, that mostly contains sequences from Sao Paulo state, probably arose on 16<sup>th</sup> October (95% HPD: 22<sup>th</sup> September – 5<sup>th</sup> November) and was defined by additional mutations NSP3:S1285F and NSP15:K12N. The other subclade that mostly comprises sequences from the North region probably arose on 29<sup>th</sup> October (95% HPD: 5<sup>th</sup> October – 17<sup>th</sup> November) and was defined by additional mutations NSP1:T170I and S:A344S (**Figure 1d**).



**Figure 1.** a) Maximum likelihood (ML) phylogenetic tree of the B.1.1.33 whole-genome sequences from Brazil. The B.1.1.33 sequences with mutation S:E484K are represented by pink (VOI N.9) and green [B.1.1.33(E484K)] circles. VOI N.9 clade is highlighted in pink. The aLRT support values are indicated in key nodes and branch lengths are drawn to scale with the left bar indicating nucleotide substitutions per site. b) Geographic distribution of the VOI N.9 identified in Brazil. Brazilian states' names follow the ISO 3166-2 standard. c) Correlation between the sampling date of B.1.1.33 sequences and their genetic distance from the root of the ML phylogenetic tree. Colors indicate the B.1.1.33 clade as indicated in a). d) Bayesian phylogeographic analysis of N.9 lineage. Tips and branches colors indicate the sampling state and the most probable inferred location of their descendent nodes, respectively, as indicated in the legend. Branch posterior probabilities are indicated in key nodes. Boxes highlight two N.9 subclades carrying additional mutations (indicated in each box). The tree was automatically rooted under the assumption of a strict molecular clock, and all horizontal branch lengths are time-scaled.

## Conclusions

In this study we identified the emergence of a new VOI (S:E484K) within lineage B.1.1.33 circulating in Brazil. The VOI N.9 displayed a low prevalence (~3%) among all Brazilian SARS-CoV-2 samples analyzed between November 2020 and February 2021, but it is already widely dispersed in the country and comprises a high fraction (35%) of the B.1.1.33 sequences detected in that period. Mutation S:E484K has been identified as one of the most important substitutions that could contribute to immune evasion as confers resistance to several monoclonal antibodies and also reduces the neutralization potency of some polyclonal sera from convalescent and vaccinated individuals<sup>11-13</sup>. Mutation S:E484K has emerged independently in multiple VOCs (P.1, B.1.351 and B.1.1.7) and VOIs (P.2 and B.1.526)<sup>14</sup> spreading around the world, and it is probably an example of convergent evolution and ongoing adaptation of the virus to the human host.

The onset date of the VOI N.9 here estimated around mid-August roughly coincides with the estimated timing of emergence of the VOI P.2 in late-July<sup>6</sup> and shortly precede the detection of a major global shift in the SARS-CoV-2 fitness landscape after October 2020<sup>15</sup>. These findings indicate that 484K variants probably arose simultaneously in the two most prevalent viral lineages circulating in Brazil around July-August, but may have only acquired some fitness advantages which accelerated its dissemination after October 2020. We predict that the Brazilian COVID-19 epidemic during 2021 will be dominated by a complex array of B.1.1.28(484K), including P.1 and P.2, and B.1.1.33(484K) variants that will completely replace the parental 484E lineages that drove the epidemic in 2020. Implementation of efficient mitigation measures in Brazil is crucial to reduce

community transmission and prevent the recurrent emergence of more transmissible variants that could further exacerbate the epidemic in the country.

The opinions expressed by the authors do not necessarily reflect the opinions of the Ministry of Health of Brazil or the institutions with which the authors are affiliated.

### **Acknowledgements**

This study was approved by the FIOCRUZ-IOC Ethics Committee (CAAE: 68118417.6.0000.5248 and 32333120.4.0000.5190) and the Brazilian Ministry of the Environment (MMA) SISGEN (A1767C3).

The authors wish to thank all the health care workers and scientists who have worked hard to deal with this pandemic threat, the GISAID team, and all the EpiCoV database's submitters, GISAID acknowledgment table containing sequences used in this study is attached to this post (**Appendix Table 2**). We also appreciate the support of the Fiocruz COVID-19 Genomic Surveillance Network (<http://www.genomahcov.fiocruz.br/>) members, the Respiratory Viruses Genomic Surveillance. General Coordination of the Laboratory Network (CGLab), Brazilian Ministry of Health (MoH), Brazilian States Central Laboratories (LACENs), Brazilian Ministry of Health (MoH), and the Amazonas surveillance teams for the partnership in the viral surveillance in Brazil. Funding support FAPEAM (PCTI-EmergeSaude/AM call 005/2020 and Rede Genômica de Vigilância em Saúde - REGESAM); Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant 403276/2020-9); Inova Fiocruz/Fundação Oswaldo Cruz (Grant VPPCB-007-FIO-18-2-30 - Geração de conhecimento).



## References:

1. Candido DS, Claro IM, de Jesus JG, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* 2020; **369**(6508): 1255-60.
2. Resende PC, Delatorre E, Gräf T, et al. Evolutionary dynamics and dissemination pattern of the SARS-CoV-2 lineage B.1.1.33 during the early pandemic phase in Brazil. *Frontier in Microbiology* 2021, **11**(February):1-14.
3. Fujino T, Nomoto H, Kutsuna S, et al. Novel SARS-CoV-2 Variant Identified in Travelers from Brazil to Japan. *Emerg Infect Dis* 2021; **27**(4).
4. Naveca F, Nascimento V, Souza V, et al. COVID-19 epidemic in the Brazilian state of Amazonas was driven by long-term persistence of endemic SARS-CoV-2 lineages and the recent emergence of the new Variant of Concern P.1. *Research Square*, 2021. <https://www.researchsquare.com/article/rs-275494/v1>
5. Faria NR, Mellan TA, Whittaker C, et al. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv* 2021.
6. Voloch CM, da Silva Francisco R, Jr., de Almeida LGP, et al. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *J Virol* 2021.
7. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* 2020; **37**(5): 1530-4.
8. Rambaut A, Holmes EC, O'Toole A, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020.
9. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2016; **2**(1): vew007.
10. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 2018; **4**(1): vey016.
11. Baum A, Fulton BO, Wloga E, et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* 2020; **369**(6506):1014-8.
12. Greaney AJ, Loes AN, Crawford KHD, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* 2021.
13. Wang P, Liu L, Iketani S, et al. Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization. *bioRxiv* 2021.
14. Gangavarapu K, Alkuzweny M, Cano M, et al. outbreak.info. Available online: <https://outbreak.info>. 2020.
15. Martin DP, Weaver S, Tegally H, et al. The emergence and ongoing convergent evolution of the N501Y lineages coincides with a major global shift in the SARS-CoV-2 selective landscape. *medRxiv* 2021.



**Appendix 1.** GISAID accession numbers of genomes lineage B.1.1.33 from this study  
EPI\_ISL\_427294 to EPI\_ISL\_427298; EPI\_ISL\_427302 to EPI\_ISL\_427304; EPI\_ISL\_456071  
to EPI\_ISL\_456077; EPI\_ISL\_456079 to EPI\_ISL\_456082; EPI\_ISL\_456084 to  
EPI\_ISL\_456087; EPI\_ISL\_456089 to EPI\_ISL\_456096; EPI\_ISL\_456098 to EPI\_ISL\_456106;  
EPI\_ISL\_467345; EPI\_ISL\_467347 to EPI\_ISL\_467353; EPI\_ISL\_467355; EPI\_ISL\_467357;  
EPI\_ISL\_467358; EPI\_ISL\_467360 to EPI\_ISL\_467365; EPI\_ISL\_467367 to EPI\_ISL\_467371;  
EPI\_ISL\_541347 to EPI\_ISL\_541350; EPI\_ISL\_541352; EPI\_ISL\_541353; EPI\_ISL\_541356 to  
EPI\_ISL\_541358; EPI\_ISL\_541360 to EPI\_ISL\_541365; EPI\_ISL\_541367; EPI\_ISL\_541369;  
EPI\_ISL\_541370; EPI\_ISL\_541376; EPI\_ISL\_541382; EPI\_ISL\_541385; EPI\_ISL\_541388;  
EPI\_ISL\_541392; EPI\_ISL\_541397 to EPI\_ISL\_541399; EPI\_ISL\_729794 to EPI\_ISL\_729798;  
EPI\_ISL\_729800; EPI\_ISL\_729802; EPI\_ISL\_729807; EPI\_ISL\_729809; EPI\_ISL\_729810;  
EPI\_ISL\_729812; EPI\_ISL\_729814; EPI\_ISL\_729816; EPI\_ISL\_729817 to EPI\_ISL\_729821;  
EPI\_ISL\_729823; EPI\_ISL\_729825; EPI\_ISL\_729826 to EPI\_ISL\_729833; EPI\_ISL\_729837 to  
EPI\_ISL\_729839; EPI\_ISL\_729841 to EPI\_ISL\_729844; EPI\_ISL\_729849; EPI\_ISL\_729851;  
EPI\_ISL\_729857; EPI\_ISL\_729858; EPI\_ISL\_729860; EPI\_ISL\_792561; EPI\_ISL\_792571 to  
EPI\_ISL\_792573; EPI\_ISL\_792579; EPI\_ISL\_792583; EPI\_ISL\_792588; EPI\_ISL\_792589;  
EPI\_ISL\_792593; EPI\_ISL\_792594; EPI\_ISL\_792596; EPI\_ISL\_792600; EPI\_ISL\_792601 to  
EPI\_ISL\_792603; EPI\_ISL\_792608; EPI\_ISL\_792610; EPI\_ISL\_792612; EPI\_ISL\_792623;  
EPI\_ISL\_792637; EPI\_ISL\_792640; EPI\_ISL\_792644; EPI\_ISL\_792648; EPI\_ISL\_1181353 to  
EPI\_ISL\_1181624.

**Appendix Table 1.** GISAID acknowledgment table.