

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

Title page

**Title:** International expansion of a novel SARS-CoV-2 mutant

**Authors:** Minjin Wang, MD

Mengjiao Li, MD

Ruotong Ren, PhD

Andreas Bråve, PhD

Sylvie van der Werf, PhD

En-Qiang Chen, PhD

Zhiyong Zong, PhD

Weimin Li, PhD

Binwu Ying, PhD

Minjin Wang, Mengjiao Li, Ruotong Ren contributed equally to this letter.

**Author Affiliations:**

Department of Laboratory Medicine, West China Hospital of Sichuan University, Chengdu, China (Minjin Wang, Mengjiao Li, Binwu Ying).

Department of Hematology, The First Hospital of Lanzhou University, Lanzhou, and Genskey Biotechnology Co., Ltd., Beijing, China (Ruotong Ren).

Department of microbiology, Public Health Agency of Sweden (Andreas Bråve).

Department of Virology, Molecular Genetics of RNA Viruses unit, CNRS UMR-3569, University of Paris, National Reference Center for Respiratory Viruses Institut Pasteur (Sylvie van der Werf).

Center of Infectious Diseases, West China Hospital of Sichuan University, Chengdu, China (En-Qiang Chen)

Department of Infection Control and Center of Infectious Diseases, West China Hospital of Sichuan University, Chengdu, China (Zhiyong Zong)

Department of Respiratory and Critical Care Medicine, West China Hospital of Sichuan University, Chengdu, China (Weimin Li)

**Corresponding author:**

Binwu Ying, PhD, Department of Laboratory Medicine, West China Hospital of Sichuan

31 University, Chengdu, China.

32 Email: [binwuying@126.com](mailto:binwuying@126.com), Tel: (0086)85423559

33 Address: Number 37 Guoxue Alley in the wuhou District, Chengdu, Sichuan, China

34

35 International expansion of a novel SARS-CoV-2 mutant

36 TO THE EDITOR:

37 SARS-CoV-2 has inevitably mutated during its pandemic spread<sup>1</sup> to cause  
38 unpredictable effects on COVID-19 and complicate epidemic control efforts. Here we  
39 report that a novel SARS-CoV-2 mutation appears to be spreading worldwide, which  
40 deserves close attention.

41 We detected 95 SARS-CoV-2 samples from Sichuan Province of China using next  
42 generation sequencing and acquired 13 whole genomes sequences, which were  
43 analyzed for sequence variation and evolution against 199 SARS-CoV-2 genomes  
44 publicly released in the GISAID EpiFlu<sup>TM</sup> database (<https://www.gisaid.org/>) and 7  
45 genomes download from NGDC database (<https://bigd.big.ac.cn/ncov>). This study was  
46 approved by the Biomedical Research Ethics Committee of West China Hospital of Sichuan  
47 University (reference no. 193, 2020) and written informed consent was obtained from all  
48 patients.

49 Based on 10 high frequency mutations (mutant allele frequency >5%), these  
50 SARS-CoV-2 genomes can be classified into 5 main groups: original stain 1 and 4  
51 variants with different mutations groups and clustering. The most common variants  
52 (Figure A, Group 1) exhibited both a missense mutation (ORF8:c.251tTa>tCa; present  
53 in 31.58% of the isolates) and a synonymous mutation (orf1ab:c.8517agC>agT; found  
54 in 30.62% of the isolates), suggesting a possible linkage between these two sites. Also,  
55 3 subgroups were evolved in the main Group 1 by other 3 mutations. Group 2 was  
56 clustered together with 3 mutants including missense variant S: c.1841gAt>gGt,  
57 orf1ab upstream gene variant and synonymous\_variant orf1ab: c.2772ttC>ttT. Group  
58 3 viral isolates were much less frequent (11.48%) and characterized by a missense  
59 mutation (orf1ab:c.10818ttG>ttT). Group 4 viral isolates contained a novel missense  
60 mutation (ORF3a:c.752gGt>gTt) first found in a Chinese family. Notably, however,  
61 Group 4 viral isolates were most frequently found outside mainland China (23.28%;  
62 27/116;  $p<0.01$  by Fisher's exact test). Additionally, Group 2 and Group 4 showed  
63 obvious aggregation in non-Chinese countries and regions.

64 The family in which the Group 4 variant was first observed in China (an older female

65 and two young family members) returned to their hometown in Sichuan from Wuhan  
66 on January 20, 2020. By January 23, the mother had a fever and cough, and her two  
67 children developed these symptoms in the following days. Nucleic acid assays  
68 performed on their throat swabs tested positive for SARS-CoV-2 on January 25. None  
69 of these individuals traveled outside of China between the start of the COVID-19  
70 epidemic and their return to Sichuan, but there the Group 4 variant first observed in  
71 this family has now demonstrated global dissemination.

72 We performed a timeline analysis using the sample collection dates reported in the  
73 GISAID EpiFlu™ database, and found that individuals infected with SARS-CoV-2  
74 strain containing the Group 4 ORF3a mutant had reached the West Coast of the  
75 United States (Orange County, California) by January 22, 2020 at the latest.  
76 Immediately afterwards, and preceding or at nearly the same time as first Group 4  
77 cases in Sichuan, additional isolates (Figure B) of this strain were reported in China  
78 (Taiwan), France (Paris), and Australia (Sydney and Clayton). According to the  
79 official records, these individuals either traveled from Wuhan, or traveled  
80 internationally prior to their disease onset. Group 4 ORF3a mutants were  
81 subsequently found in several other countries, including Singapore, South Korea, the  
82 United Kingdom and Italy. It should be noted that this mutant virus strain appears to  
83 be the most prevalent form of SARS-COV-2 in France, Italy, Brazil, and Singapore  
84 (Figure C).

85 Virus genome data from France indicate that SARS-CoV-2 strains carrying  
86 ORF3a:c.752gGt>gTt often have a S:c.1099Gtc>Ttc mutation in their S gene, which  
87 interacts with ACE2 to mediate viral entry into its host cells<sup>3</sup>, and is regarded as a  
88 critical factor for viral transmission and virulence<sup>4, 5</sup>. It is not clear if this mutation  
89 enhances host cell entry but this information would be of great importance in  
90 assessing the potential for increased virulence of Group 4 SARS-CoV-2 strains  
91 carrying this mutation. It is also not known how common this mutation is in Group 4  
92 viral isolates from different geographical regions. Given the prevalence of Group 4  
93 isolates in multiple countries, including France, Italy and South Korea, which is  
94 experiencing a rapidly growing epidemic, this information should be of significant

95 interest.

96 At present, the SARS-CoV-2 epidemic in China seems to be diminishing in response  
97 to control efforts, but the rapid global spread of this new virus, and its mutants, has  
98 become a major health concern. Very little is known about how rapidly the  
99 SARS-CoV-2 genome mutates and how this affects transmission or disease severity.  
100 Better understanding of these factors should be useful in efforts to curtail the global  
101 and regional spread of this virus.

102

103 **Conflict of Interest Disclosures:** None reported.

104 **Funding/Support:** This study was funded by Science & Technology Department of Sichuan  
105 Province (Grant number: 2020YFS0004) and West China Hospital of Sichuan University  
106 (Grant number: HX-2019-nCoV-066).

107 **Additional Contributions:** We thank Dr Wan Xiong for polishing the manuscript language.

108 We also thank Shuo Guo and Yanbing Zhou for collecting literature.

109

110 **Reference :**

- 111 1. [www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/](http://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/)
- 112 2. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data -  
113 from vision to reality[J]. Euro surveillance : bulletin Europeen sur les maladies  
114 transmissibles = European communicable disease bulletin. 2017,22(13).  
115 doi:org/10.2807/1560-7917.es.2017.22.13.30494
- 116 3. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor  
117 usage for SARS-CoV-2 and other lineage B betacoronaviruses[J]. Nature  
118 microbiology. 2020. doi:org/10.1038/s41564-020-0688-y
- 119 4. Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining 'host jump' of  
120 coronaviruses SARS-CoV, MERS-CoV, and beyond[J]. Trends in microbiology.  
121 2015,23(8):468-478. doi:org/10.1016/j.tim.2015.06.003
- 122 5. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H. Tissue

123 distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first  
124 step in understanding SARS pathogenesis[J]. The Journal of pathology.  
125 2004,203(2):631-637. doi:org/10.1002/path.1570

126

127

128 **Figure legend :**

129 **Figure: Maximum likelihood tree based on the whole genome sequences of 221**  
130 **viral strains.**

131 A) 199 high quality genomes were collected from GISAID EpiFlu™ database,  
132 including 1 *Rhinolophus affinis* isolate, 6 *Manis javanica* isolates and 2 environmental  
133 isolates. 22 additional genomes were collected from other resource, including 7  
134 genomes from NGDC (<https://bigd.big.ac.cn/ncov>), 13 genomes from WCH.  
135 SARS-CoV (NC\_004718.3) and MERS-CoV (NC\_019843.3) genomes sequence  
136 were downloaded from NCBI RefSeq database. MAFFT (version 7.543) was used for  
137 sequence alignment, and PhyML (version 3.0) was used to construct the evolutionary  
138 tree. Variation information of human SARS-CoV-2 genome was derived from NGDC.  
139 Mutations of 13 WCH genomes were analyzed using NGDC online tools  
140 (<https://bigd.big.ac.cn/ncov/tool/variation-identify>).

141 B) Location and collection time of ORF3a:c.752gGt>gTt variant genomes.

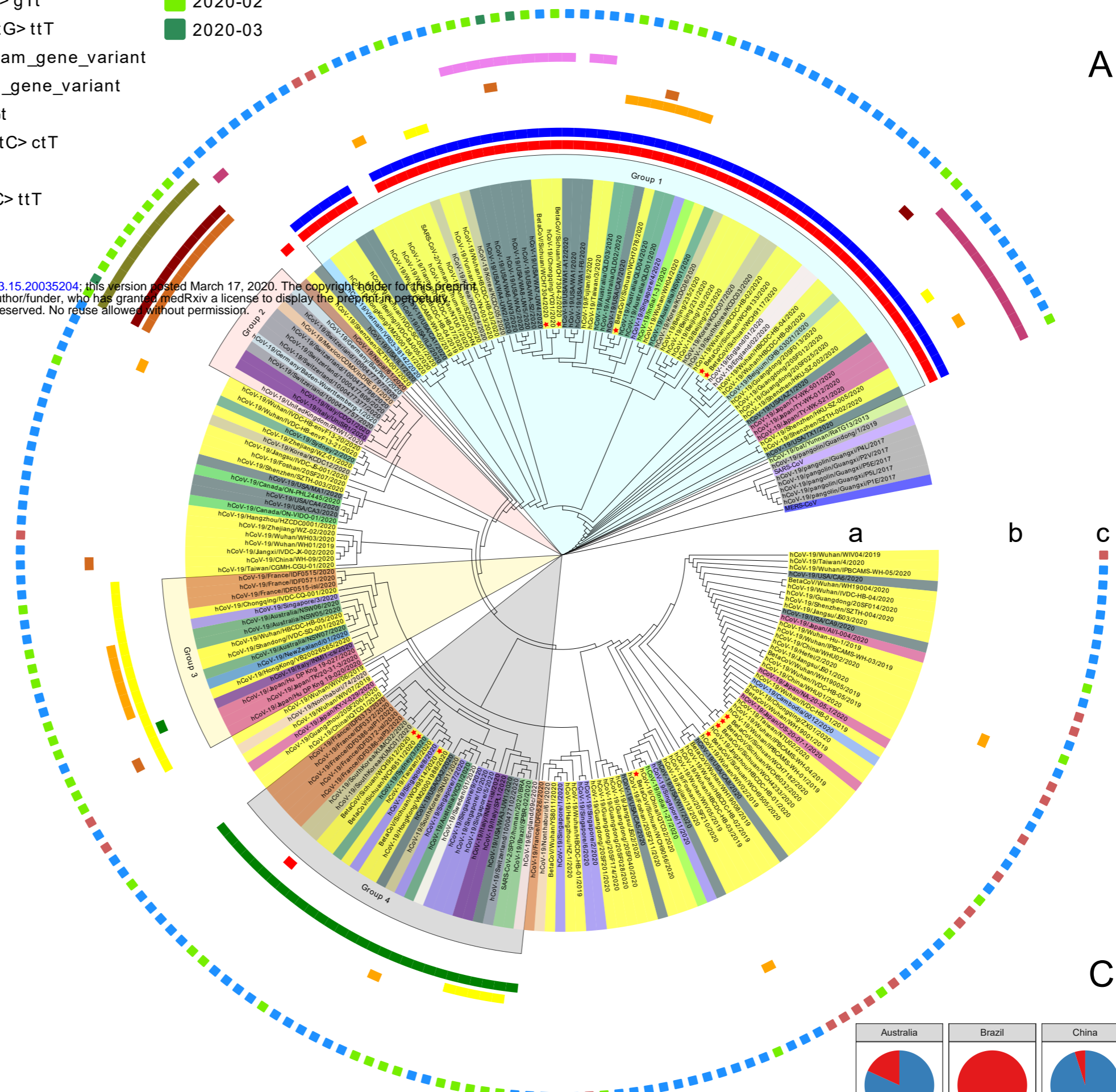
142 C) Composition of variant and non-variant genomes of ORF3a:c.752gGt>gTt in  
143 different countries.

- Group**
- Australia
  - Bat
  - Belgium
  - Brazil
  - Cambodia
  - Canada
  - China
  - France
  - Germany
  - India
  - Italy
  - Japan
  - MERS
  - Mexico
  - Nepal
  - New Zealand
  - Pangolin
  - SARS
  - Singapore
  - South Korea
  - Sweden
  - Switzerland
  - Thailand
  - USA
  - United Kingdom
  - Vietnam

- Variation**
- ORF8:c.251tTa> tCa
  - orf1ab:c.8517agC> agT
  - ORF3a:c.752gGt> gTt
  - orf1ab:c.10818ttG> ttT
  - ORF10:downstream\_gene\_variant
  - orf1ab:upstream\_gene\_variant
  - S:c.1841gAt> gGt
  - orf1ab:c.17796ctC> ctT
  - N:c.822ttC> ttT
  - orf1ab:c.2772ttC> ttT

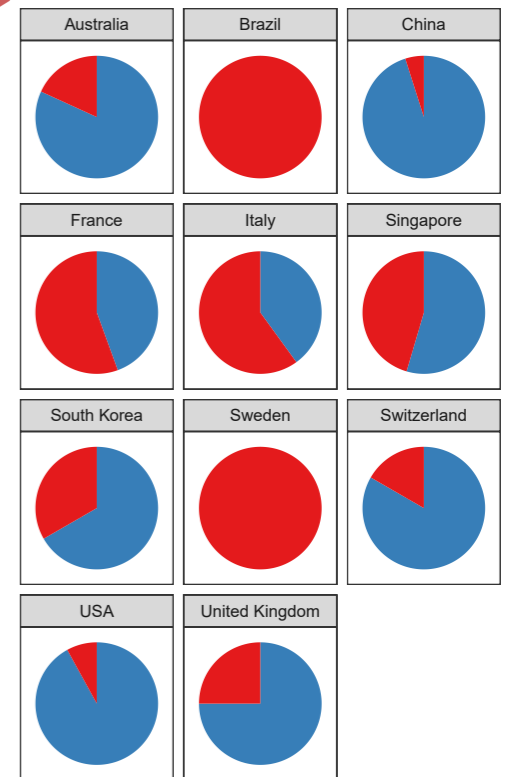
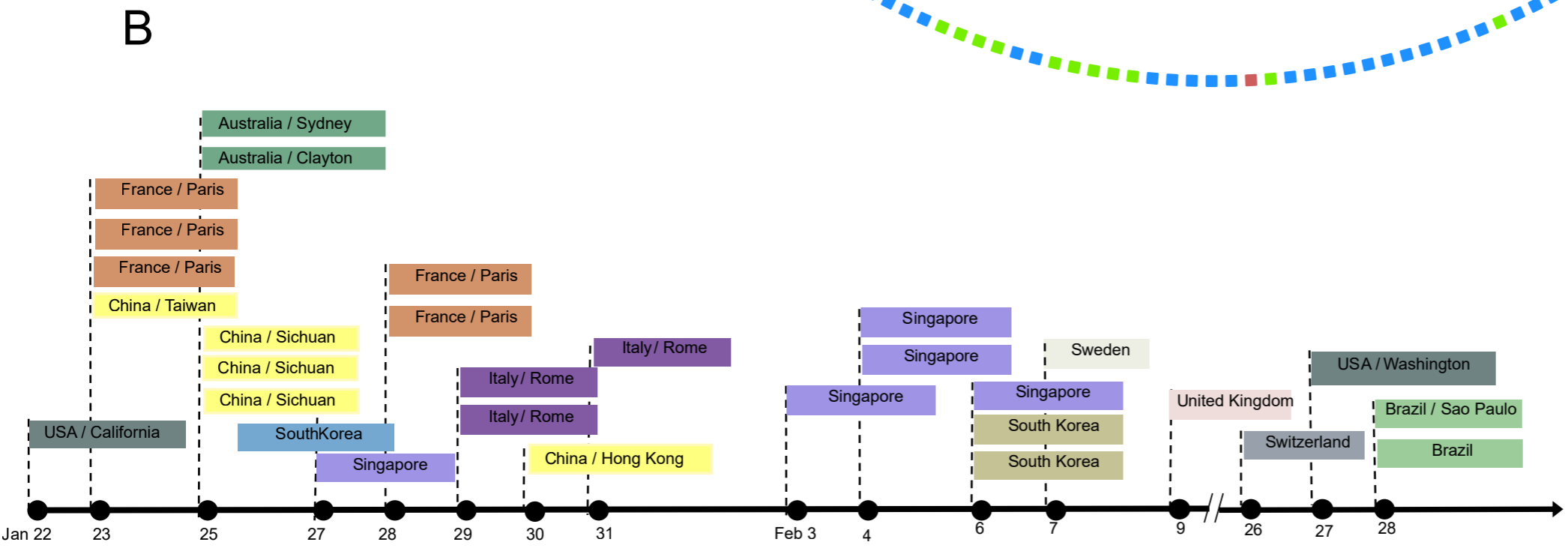
- Time**
- 2019-12
  - 2020-01
  - 2020-02
  - 2020-03

medRxiv preprint doi: <https://doi.org/10.1101/2020.03.15.20035204>; this version posted March 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



A

C



Type  
■ Variant  
■ non\_Variant