

1 **Transmission of B.1.617.2 Delta Variant between vaccinated healthcare workers**

2

3 Steven A. Kemp<sup>1,2†</sup>, Mark TK Cheng<sup>1†</sup>, Will Hamilton<sup>2</sup>, Kimia Kamelian<sup>3</sup>, INSACOG Consortium, Sujit  
4 Singh,<sup>4</sup> Partha Rakshit<sup>4</sup>, Anurag Aggrawal<sup>5</sup>, Chris Illingworth<sup>6,7,8†</sup> & Ravindra K. Gupta<sup>1,2,9†\*</sup>

5

6 1. Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Cambridge,  
7 UK.

8 2. Department of Medicine, University of Cambridge, Cambridge, UK.

9 3. Public Health Agency of Canada, National Microbiology Laboratory, Winnipeg, Manitoba,  
10 Canada

11 4. National Centre for Disease Control, Delhi, India

12 5. CSIR Institute of Genomics and Integrative Biology, Delhi, India

13 6. MRC-University of Glasgow Centre for Virus Research, Garscube Campus, 464 Bearsden  
14 Road, Glasgow, UK

15 7. MRC Biostatistics Unit, University of Cambridge, East Forvie Building, Forvie Site, Robinson  
16 Way, Cambridge, UK

17 8. Department of Applied Mathematics and Theoretical Physics, University of Cambridge,  
18 Cambridge, UK

19 9. Africa Health Research Institute, Durban, South Africa.

20

21 \*Corresponding Author

22 Ravindra K Gupta: Jeffrey Cheah Biomedical Centre, Cambridge, CB5 8UB, United Kingdom

23 †Authors contributed equally

24

25

26

27 **Abstract**

28 Breakthrough infections with SARS-CoV-2 Delta variant have been reported in vaccine recipients and  
29 in individuals infected with previous variants. However the potential for fully vaccinated individuals  
30 (two doses) to transmit SARS-CoV-2 is unclear. We here analyse data from health care workers in  
31 two hospitals in India, constructing probable transmission networks from epidemiological and virus  
32 genome sequence data using a suite of computational approaches. Among known cases we identify  
33 a high probability that doubly vaccinated individuals transmitted SARS-CoV-2, and potential cases of  
34 virus transmission between individuals who had received two doses of vaccine. Our findings  
35 highlight the need for ongoing infection control measures even in highly vaccinated populations.

36

37

## 38 Introduction

39 The B.1.617.2 (Delta) variant is now the most common lineage circulating in India and worldwide.  
40 Vaccine efficacy (VE) for the Delta variant is estimated at around 50% for mild symptomatic infection  
41 after two doses of either mRNA or adenovirus-based vaccine platforms<sup>1</sup>. However, VE for  
42 moderately-severe and severe disease is estimated as 95%<sup>1</sup>. Breakthrough infections of the Delta  
43 variant have been reported in vaccine recipients, as well as re-infection of individuals infected with  
44 previously circulating variants<sup>2-4</sup>. We previously reported 155 occurrences of vaccine breakthrough  
45 infections<sup>5</sup> amongst frontline Delhi healthcare workers (HCW), all of whom had one or two doses of  
46 the Covishield (ChAdOx1) vaccine, the majority of whom were vaccinated >21 days prior to  
47 symptomatic presentation.

48

49 Existing literature suggests that there are similar viral loads between vaccinated and unvaccinated<sup>6,7</sup>,  
50 but with steeper decay<sup>8</sup> over time in vaccinated persons. Furthermore the probability of culturing  
51 virus appears lower in the vaccinated. It is unclear whether individuals with vaccine breakthrough  
52 are able to transmit to other vaccinated individuals, or whether transmission is driven solely by the  
53 unvaccinated. Here we utilise *in silico* approaches combining genomic and epidemiological  
54 investigation data to identify probable Delta variant transmission events between vaccinated HCW in  
55 Indian hospitals.

56

## 57 Results

### 58 ChAdOx1 nCoV-19 Breakthrough Infection

59 All symptomatic HCWs in two hospitals who had received one or two doses of the ChAdOx1 nCoV-19  
60 mRNA vaccine (AZD1222), in addition to a set of unvaccinated HCWs, underwent RT-PCR testing  
61 within one day of symptom onset as part of a hospital staff symptomatic testing program. From  
62 these cases we identified cases of breakthrough infection in staff who had received two doses of the  
63 vaccine. In hospital A, there were 81 breakthrough infections amongst 1100 HCWs, and in hospital B,  
64 32 infections amongst 4000 HCWs, as previously reported<sup>5</sup>.

65

66 Among the 113 cases from hospitals A and B, 12.4% were administrative staff, 31.9% were nurses,  
67 40.7% were primary physicians, 7.1% were paramedics and 3.5% were pharmacists. The remaining  
68 3.5% of breakthrough infections consisted of medical interns (1.8%), and laboratory workers (1.8%)  
69 (**Supplementary Table 1**). The median interval between receiving a second vaccination and date of  
70 positive RT-PCR test was 45 days (range 3-78 days). Amongst the breakthrough infections of doubly-  
71 vaccinated HCW, 90.7% were infected with B.1.617.2, 5.3% by B.1.1.7-like, 1.3% by B.1.538.

72 **(Supplementary Table 2)**. The most commonly reported symptoms were fever (82.3% of all cases),  
73 cough (43.4%), myalgia (20.4%), and loss of smell/taste (14.2%) **(Supplementary Table 3)**.

74

75 Whole genome sequencing was used to characterise staff nose and throat swab samples. In hospital  
76 A, 66 sequences with high quality whole genome coverage of >90% were generated, including 43  
77 cases of breakthrough infection. In hospital B, high quality genome sequences were generated from  
78 52 samples, including all 32 symptomatic breakthrough infections.

79

### 80 **Community Sequencing**

81 Between April – May 2021, >99% of SARS-CoV-2 sequences in India were assigned Delta (B.1.616.2)  
82 lineages. However, since August 2021, the Delta sub-lineages have proliferated - predominantly  
83 AY.12 – across India and elsewhere<sup>9</sup>. To determine the relationship between community and HCW  
84 sequences, we inferred a maximum-likelihood phylogeny to estimate dispersion **(Supplementary**  
85 **Figure 1)**. The analysis suggested multiple introductions into hospitals A and B, with subsequent  
86 intra-hospital transmission. We found significant partitioning of the inferred phylogeny, separating  
87 the sequences into distinct clades. Mutations relative to the Delta consensus that were found in  
88 these cases were spread across the SARS-CoV-2 genome. None of the identified SNPs were in  
89 homoplasic regions of the genome. We note that the individuals 115, 127, and 305 in hospital A had  
90 identical virus genomes.

91

### 92 **Linkage between vaccinated HCW**

93 Considering cases of infection with the B.1.617.2 Delta variant, we inferred putative transmission  
94 pairs using the A2B-COVID<sup>10</sup> software package, which identifies pairs of individuals for whom  
95 symptom onset dates and viral genome sequences are consistent with direct transmission. Using  
96 parameters inferred for the Delta SARS-CoV-2 lineage **(Supplementary Text, Supplementary Figure**  
97 **2)** we identified 35 putative transmission events involving 14 HCWs in hospital A, and 26 putative  
98 events involving 13 patients in hospital B **(Supplementary Figures 3, 4)**.

99

100 Data from these cases were then analysed with the A2B-Network software package<sup>11</sup> which  
101 calculates the probability of specific networks of transmission events between a set of individuals.  
102 To reduce excessive computational load, potential transmission events in hospital A were filtered to  
103 remove those involving the gain of more than one SNP. From the remaining cases we identified a  
104 network involving six HCWs, two of whom had received their second dose of vaccine at least 14 days  
105 prior to reporting symptoms **(Figure 1)**. A total of 1381 possible transmission networks between

106 these individuals were identified, allowing us to calculate the probabilities of different sets of  
107 transmission events. Under the conditions of our model there was a 92.1% chance that one of the  
108 two individuals receiving a second vaccine 14 days prior to infection was the source of transmission  
109 to another individual in the network. If this criterion is relaxed to 7 days between second  
110 vaccination and infection, we estimated a 99.9% chance that at least one of the three such  
111 individuals infected another individual in the network, with a 98.7% probability of direct virus  
112 transmission between these three individuals. The filtering of this subset individuals from the larger  
113 network could not have increased any of these statistics (**Supplementary Figure 5**). Other unrelated  
114 networks in hospital A did not involve transmission from vaccinated HCWs to unvaccinated HCWs  
115 (**Supplementary Figure 6**).

116

117 Analysis of potential transmission events in hospital B identified a network involving eight HCWs,  
118 four of whom had received their second dose of vaccine at least 14 days prior to reporting  
119 symptoms (**Figure 1**). A total of 128 possible transmission networks between these individuals were  
120 identified, all of which implied that one of the four individuals who received a second vaccine at  
121 least 14 days prior to infection was the source of transmission to another individual in the network.  
122 Among the five individuals who were infected more than seven days post second vaccination, we  
123 found a 97.8% chance of a transmission event having occurred between individuals vaccinated with  
124 two doses.

125

126 While our analysis could not account for unobserved data, we performed validation calculations  
127 evaluating the self-consistency of the observed sequence data with the inferred patterns of virus  
128 transmission. Comparing our data against simulations suggested that the transmission network  
129 inferred for hospital A is self-consistent with the observed sequence data, though the network  
130 inferred for hospital B is potentially distorted by missing data, with too high a number of individuals  
131 carrying unique mutations (**Supplementary Figure 7**). Details of simulations are provided in  
132 **Supplementary Text**.

133

134

### 135 **Transmission Networks**

136 Inferred transmission networks involved a variety of medical staff. In hospital A, a common genome  
137 sequence was shared by the junior medical staff member P115, the nursing student P127, and by  
138 P305, a member of the nursing staff, who collectively became symptomatic over a period of six days.  
139 Three other staff had virus genomes which differed from this by a single SNP each. The

140 ophthalmologic junior medical staff P142 became symptomatic a day after P305; the SNP 8016T  
141 causes a nonsynonymous change in ORF1ab, A2584V. P155 was also a member of the nursing staff;  
142 the SNP C9509T leads to the nonsynonymous substitution T3082I in ORF1ab. Finally, P164 was a  
143 paramedic; the SNP A19298G represents a nonsynonymous change in the ORF1ab (Y6345C).

144

145 In hospital B, the individuals P205, P207, and P240 all received their second dose of vaccine at least  
146 14 days before testing positive. P205 was a paediatrician, P207 and P240 both adult physicians.  
147 Further details of symptoms, job titles, symptoms, and positive test dates, where known, are shown  
148 in **Supplementary Table 4**. Full details of SNPs are given in **Supplementary Figure 8 and 9**.

149

## 150 **Discussion**

151 As the Delta variant dominates globally, information on the transmission amongst fully vaccinated  
152 individuals is needed to dictate the need for future infection control and vaccination boost  
153 strategies. Our study highlights the plausibility of Delta variant transmission in a healthcare setting  
154 from and between fully vaccinated individuals. Our study has caveats; firstly by only considering  
155 symptomatic cases, we exclude HCWs who are asymptomatic or fail to report their infection status.  
156 Secondly, the number of valid samples is further reduced by incomplete sampling and subsequent  
157 sequence filtering, likely resulting in incomplete transmission chains. Calculations suggested that the  
158 networks inferred for hospital B are likely to be affected by missing data, though the data from  
159 hospital A was not clearly affected.

160

161 Although vaccination is still highly effective for Delta in protecting against severe disease<sup>12</sup>, we have  
162 demonstrated that breakthrough infections still occur in healthcare settings in individuals within 60  
163 days of the second dose when circulating neutralising antibody levels are at their highest<sup>13</sup>. Given  
164 the risk of onward transmission to potentially very vulnerable patients (including people in whom  
165 vaccination is less effective, such as those with compromised immune systems), our findings  
166 highlight the need for ongoing infection control measures even in highly vaccinated populations, and  
167 mask wearing in individuals who have received two doses of vaccine, in order to limit onward  
168 transmission.

169

170

171 **References**

- 172 1. Pramod S, Govindan D, Ramasubramani P, et al. Effectiveness of Covishield vaccine in  
173 preventing Covid-19 – A test-negative case-control study. medRxiv 2021:2021.07.19.21260693.
- 174 2. Jolly B, John N, Bhojar RC, et al. Genomic survey of SARS-CoV-2 vaccine breakthrough  
175 infections in healthcare workers from Kerala, India. Journal of Infection 2021;83:237-79.
- 176 3. Tyagi K, Ghosh A, Nair D, et al. Breakthrough COVID19 infections after vaccinations in  
177 healthcare and other workers in a chronic care medical facility in New Delhi, India. Diabetes &  
178 Metabolic Syndrome: Clinical Research & Reviews 2021;15:1007-8.
- 179 4. Dhar MS, Marwal R, VS R, et al. Genomic characterization and epidemiology of an emerging  
180 SARS-CoV-2 variant in Delhi, India. Science;0:eabj9932.
- 181 5. Mlcochova P, Kemp S, Dhar MS, et al. SARS-CoV-2 B.1.617.2 Delta variant replication and  
182 immune evasion. Nature 2021.
- 183 6. Riemersma KK, Grogan BE, Kita-Yarbro A, et al. Vaccinated and unvaccinated individuals  
184 have similar viral loads in communities with a high prevalence of the SARS-CoV-2 delta variant.  
185 medRxiv 2021:2021.07.31.21261387.
- 186 7. Pouwels KB, Pritchard E, Matthews PC, et al. Effect of Delta variant on viral burden and  
187 vaccine effectiveness against new SARS-CoV-2 infections in the UK. Nature Medicine 2021.
- 188 8. Chia PY, Xiang Ong SW, Chiew CJ, et al. Virological and serological kinetics of SARS-CoV-2  
189 Delta variant vaccine-breakthrough infections: a multi-center cohort study. medRxiv  
190 2021:2021.07.28.21261295.
- 191 9. INSACOG. INSACOG WEEKLY BULLETIN2021 23/08/21.
- 192 10. Illingworth CJR, Hamilton WL, Jackson C, et al. A2B-COVID: A method for evaluating potential  
193 SARS-CoV-2 transmission events. medRxiv 2021:2020.10.26.20219642.
- 194 11. Illingworth CJR, Hamilton WL, Warne B, et al. Superspreaders drive the largest outbreaks of  
195 hospital onset COVID-19 infections. Elife 2021;10:e67308.
- 196 12. Tartof SY, Slezak JM, Fischer H, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up  
197 to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet  
198 2021;398:1407-16.
- 199 13. Wall EC, Wu M, Harvey R, et al. Neutralising antibody activity against SARS-CoV-2 VOCs  
200 B.1.617.2 and B.1.351 by BNT162b2 vaccination. Lancet 2021;397:2331-3.
- 201 14. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:  
202 improvements in performance and usability. Mol Biol Evol 2013;30:772-80.
- 203 15. Rambaut A, Holmes EC, O'Toole A, et al. A dynamic nomenclature proposal for SARS-CoV-2  
204 lineages to assist genomic epidemiology. Nat Microbiol 2020;5:1403-7.

205 16. Page AJ, Taylor B, Delaney AJ, et al. SNP-sites: rapid efficient extraction of SNPs from multi-  
206 FASTA alignments. *Microbial Genomics* 2016;2.

207 17. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of  
208 single nucleotide polymorphisms, SnpEff. *Fly* 2012;6:80-92.

209

210

## 211 **Methods**

### 212 **Study Design**

213 From 25th March – 10th May 2021 all frontline healthcare workers who presented with symptoms  
214 consistent with SARS-CoV-2 were diagnostically tested for the presence of SARS-CoV-2 by means of  
215 reverse transcriptase polymerase chain reaction (RT-PCR) using TRUENAT or CBNAAT (GENEXPERT).  
216 Findings were expressed as the cycle threshold (Ct) for the gene encoding the nucleocapsid protein  
217 (N gene) for hospital A, and the gene encoding for the envelope protein (E gene) for hospital V. A Ct  
218 value of less than 30 was considered to be infective. A vaccine breakthrough infection was defined  
219 as detection of SARS-CoV-2 in a sample collected from an individual 21 days after receipt of a second  
220 dose of ChAdOx1.

221

222 Ethical approval for the study of vaccine elicited antibodies in sera from vaccinees was obtained  
223 from the East of England Cambridge Central Research Ethics Committee Cambridge (REC ref.  
224 17/EE/0025). Use of convalescent sera had ethical approval from the South Central Berkshire B  
225 Research Ethics Committee (REC ref. 20/SC/0206; IRAS 283805). Studies involving HCWs (including  
226 testing and sequencing of respiratory samples) were reviewed and approved by The Institutional  
227 Human Ethics Committees of the National Centre for Disease Control (NCDC) and Council Of  
228 Scientific And Industrial Research–Institute Of Genomics And Integrative Biology (CSIR IGIB)  
229 (NCDC/2020/NERC/14 and CSIRIGIB/IHEC/202021/01). Participants provided informed consent.

230

### 231 **Bioinformatics and Phylogenetic analysis**

232 Fasta consensus sequences were obtained from two separate Hospitals in Delhi, India. All sequences  
233 were concatenated into a single fasta and aligned to reference strain MN908947.3 with mafft  
234 v7.847<sup>14</sup> using the --keeplength and --addfragments options. Following this, all sequences were  
235 screened for number of gaps and N-regions using the Nextclade v1.5.4  
236 (<https://clades.nextstrain.org/>) server. All sequences were assigned a lineage with Pangolin  
237 v3.1.11<sup>15</sup>, pangoleARN (dated 9<sup>th</sup> August 2021) and scorpio v0.0.14. Sequences that could not be



238 assigned a lineage were discarded. After assigning lineages, all sequences with more than 5% N-  
239 regions were also discarded.

240

241 To contextualise outbreak sequences, all sequences from India with lineage defined as B.1.617.2  
242 from the month of April 2021 were downloaded from the Global Initiative on Sharing Avian Influenza  
243 Data (GISAID) EpiCoV database. Incomplete (<29,000 base pair), duplicate, and low-quality  
244 sequences (defined as equal to or more than 5% Ns; less than 95% genome coverage) were  
245 excluded.

246

#### 247 **SNP distance, variant calling and annotation**

248 Single-nucleotide polymorphisms (SNPs), relative to the Delta variant strain, were identified by re-  
249 alignment of groupings of sequences based on each hospital to the Delta variant consensus  
250 reference (MZ359841.1) using mafft v7.8477 with the --keeplength and --addfragments parameters.  
251 Initial analysis of pairwise SNP distance between each patient was conducted using snp-dists v0.8.2  
252 with default parameters.

253

254 Following this, a VCF of acquired mutations of each patient with respect to the reference strain  
255 MN908947.3 (Wuhan-Hu-1) is calculated by snp-sites v2.5.1<sup>16</sup> using the -v and -c option. Multi-allelic  
256 variants are broken down into biallelic variant representations and are subsequently annotated by  
257 snpEff v5.0e<sup>17</sup> with reference to MN908947.3. The nucleotide and amino acid variants between each  
258 transmission pairs are extracted from the VCF, and SNP number verified using an in-house script  
259 ([https://github.com/TKMarkCheng/Indian\\_HCW](https://github.com/TKMarkCheng/Indian_HCW)). The VCFs were manually scanned to identify  
260 homoplasic/problematic sites ([https://github.com/W-L/ProblematicSites\\_SARS-CoV2](https://github.com/W-L/ProblematicSites_SARS-CoV2)).

261

#### 262 **Estimating likelihood of person-to-person transmission**

263 As an initial assessment of whether participants in the study had passed an infection to another, we  
264 utilised A2B-COVID<sup>10</sup>. This software considers data from individuals in a pairwise fashion, considering  
265 the timing of symptom onset and virus genome sequences in order to assess for each pair of  
266 individuals A and B whether the data are consistent with an underlying model of direct virus  
267 transmission from A to B. Data from each pair are described as 'consistent' with transmission,  
268 'borderline', or 'unlikely' to have been produced given direct transmission.

269

270 Network reconstruction was performed using the A2B-Network software<sup>11</sup>. Again using dates of  
271 symptom onset and virus genome sequences, this identifies plausible networks via which all of the

272 individuals in a set could have transmitted the virus between themselves, and assesses the  
273 probability of each such network having occurred, given our model assumptions. In this manner, the  
274 code produces ensembles of networks, describing the extent to which the data constrain the  
275 possible routes of transmission of the virus. The probabilities we report were calculated by summing  
276 network likelihoods. For example, the probability that a network contains a transmission event  
277 between doubly vaccinated individuals is the sum of the probabilities of the networks which contain  
278 at least one such transmission event. Our model assumes a set of underlying parameters describing  
279 the transmission dynamics and rate of evolution of the SARS-CoV-2 virus. The network probabilities  
280 we report are conditional on these parameters and upon the assumption that the individuals to  
281 whom we apply the model are connected by a transmission network.

282

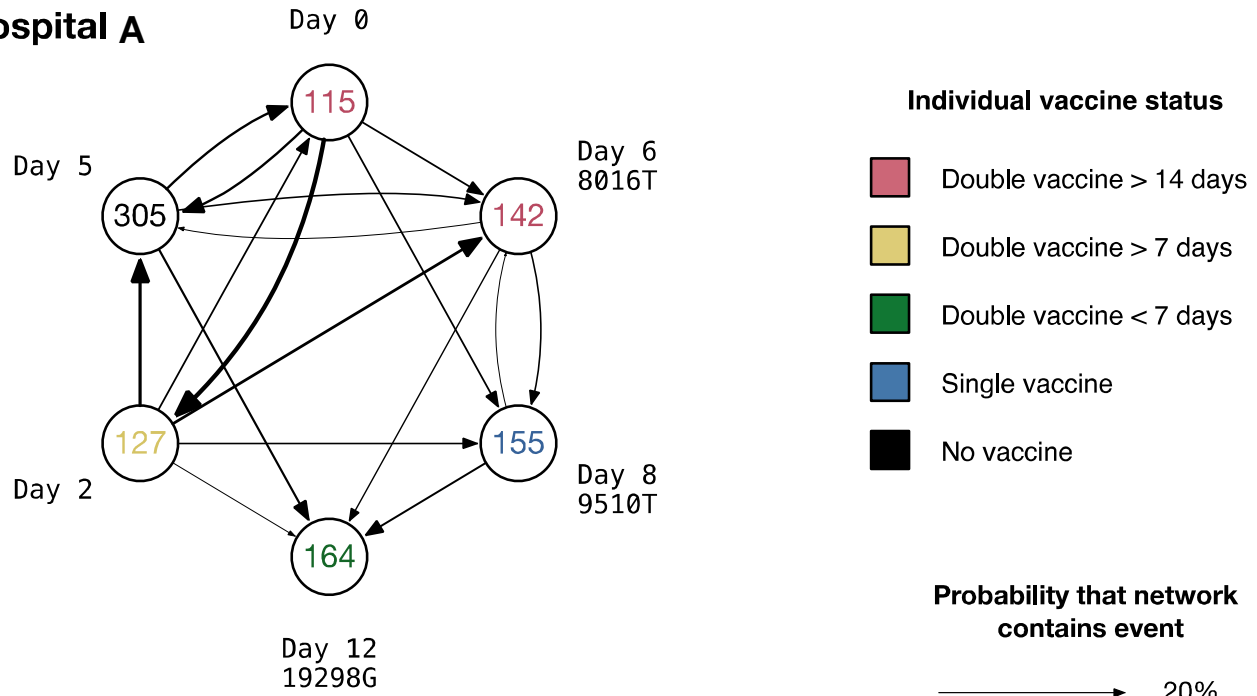
### 283 **Network validation**

284 Transmission networks were validated for self-consistency using a model of simulated SARS-CoV-2  
285 outbreaks. Simulations used identical parameters to the network inference model to describe the  
286 transmission dynamics and rate of evolution of the SARS-CoV-2 virus. Via repeated simulation, we  
287 evaluated whether the genetic properties of the virus genome sequences in our datasets, described  
288 in terms of the number of unique mutations observed, and the number of mutations unique to each  
289 individual, were consistent with what we would expect to observe if the networks we had inferred  
290 were complete and correct. Further details are given in **Supplementary Text**.

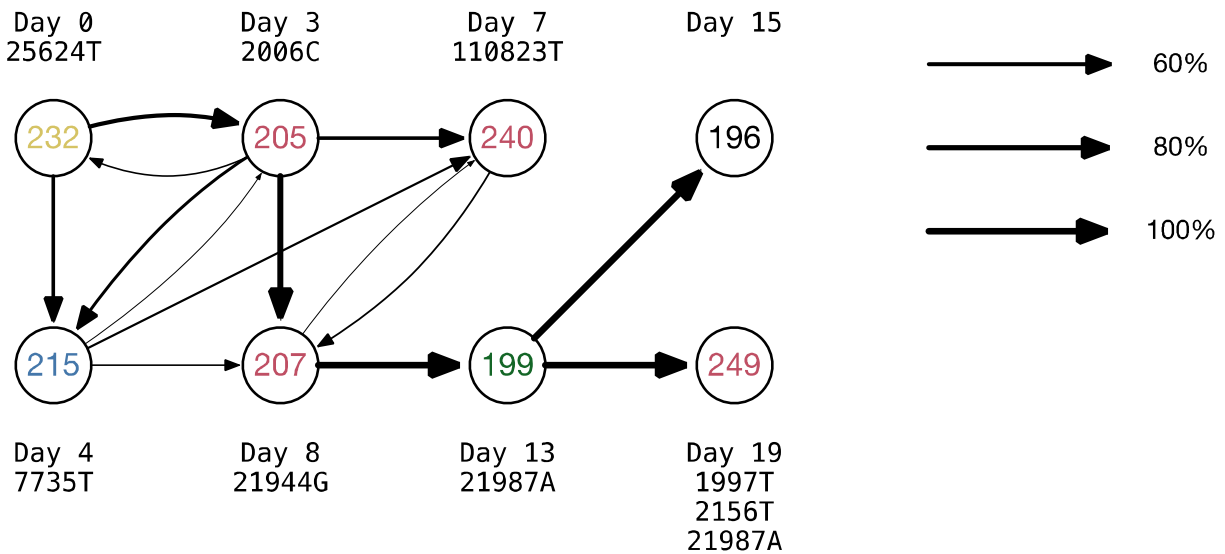
291

It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

## Hospital A



## Hospital B



**Figure 1. Potential transmission networks between HCWs.** Individual labels are coloured according to vaccine status, including the timing prior to infection at which the second vaccine was given, where relevant. The thickness of lines between individuals show the probabilities of distinct pairwise transmission events between individuals; these probabilities are conditional on transmission having occurred between the individuals observed in each network. Labels show the relative dates on which individuals became symptomatic, and respective gains of nucleotides in sequences collected from each individual with respect to the mutual consensus.