



Figure 4—figure supplement 2. Phylogenetic tree of 34 healthcare worker (HCW) SARS-CoV-2 genomes. Branch tips are coloured by HCW base ward. 34/35 sequenced genomes passed the filter of <2990 (~10%) N. A SARS-CoV-2 genome collected in Wuhan in December 2019 was selected to root the tree, visualised initially on Nextstrain (<https://nextstrain.org/>) and the fasta file was downloaded from GISAID (ID: EPI_ISL_402123) (<https://www.gisaid.org/>). Multiple sequence alignment of consensus fasta files was performed using MAFFT with default settings (Katoh K. MAFFT version 7. <https://mafft.cbrc.jp/alignment/software/>). The alignment was manually inspected using AliView (University U. AliView. <https://ormbunkar.se/aliview/>). A maximum likelihood tree was produced using IQ-TREE software (<http://www.iqtree.org/>) with ModelFinder Plus option (-m MFP), which chooses the nucleotide substitution model that minimises Bayesian information criterion (BIC) score. The model "chosen" was TPM2u+F (details: <http://www.iqtree.org/doc/Substitution-Models>). The tree was manually inspected in FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>), rooted on the 2019 Wuhan sample, ordered by descending node and exported as a Newick file. The tree was visualised in the online software Microreact (<https://microreact.org/showcase>) in a private account, exported as a png image, which is shown here. Due to low genetic diversity in the virus (very recent introduction) genomic similarity alone cannot be used to infer transmission chains, as viruses can be identical by chance. Achieving higher resolution on transmission chains requires integrating clinical and detailed epidemiological data with genomic data from HCW and patients to uncover plausible transmission pathways.