

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Confirmed  |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested   |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	NA
Data analysis	<p>See Supplementary Data for settings used for all tools</p> <p>Minimap2 v2.17</p> <p>SAMtools v1.10</p> <p>iVar v1.2.1</p> <p>MASH v2.2.2</p> <p>Quicktree v2.5</p> <p>MAFFT v7.453</p> <p>Arbow v0.4.0 (<a href="https://github.com/MDU-PHL/arbow">https://github.com/MDU-PHL/arbow</a>)</p> <p>IQ-Tree v1.6.12</p> <p>FigTree v1.4.4</p> <p>ClusterPicker v1.2.3</p> <p>Pangolin tool v1.1.11 (<a href="https://github.com/hCoV-2019/pangolin">https://github.com/hCoV-2019/pangolin</a>), pangolin lineages 2020-05-07 (database)</p> <p>TempEst 1.5.3</p> <p>BEAST v2.5</p> <p>blast v2.10.0+</p> <p>R for statistical computing v3.5.2</p>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Consensus sequences and Illumina sequencing reads were deposited into GenBank under BioProject PRJNA613958 (Supplementary Data); a small subset were rejected by NCBI, which we are currently investigating. Consensus genome sequences and multiple-sequence alignment for ALL samples are also available from <https://github.com/MDU-PHL/COVID19-paper>.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	To present a accurate understanding of the genomics of SARS-CoV-2 in our population every available samples underwent genome sequencing and was included in the study.
Data exclusions	To ensure high quality analysis, we applied quality control criteria for the sequence data. We applied quality control checks on consensus sequences, requiring $\geq 80\%$ genome recovered, $\leq 25$ single nucleotide polymorphisms (SNPs) from the reference genome, and $\leq 300$ ambiguous or missing bases for sequences to 'pass' QC. Duplicate sequences from the same patient were excluded for the multiple sequence alignment, but used in the intra-patient variability analysis.
Replication	Ten samples were sequenced in duplicate in two different sequencing runs, with zero SNP differences found between the two sequences for all ten samples.
Randomization	Randomisation was not relevant to this study as all samples were subjected to the same genomic analyses in order to fully characterise the dataset.
Blinding	Scientists performing phylogenetic and phylodynamic analyses were blinded to the patient and epidemiologic data, such that genomic cluster determination was performed without knowledge of epidemiologic clustering.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- | n/a                                 | Involvement in the study                               |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |

### Methods

- | n/a                                 | Involvement in the study                        |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |