









Genomic assessment of quarantine measures to prevent SARS-CoV-2 importation and transmission

Dinesh Aggarwal ^{1,2,3,4}✉, Andrew J. Page ^{5,104}, Ulf Schaefer^{2,104}, George M. Savva⁵, Richard Myers², Erik Volz ⁶, Nicholas Ellaby², Steven Platt², Natalie Groves², Eileen Gallagher², Niamh M. Tumelty ⁷, Thanh Le Viet ⁵, Gareth J. Hughes ⁸, Cong Chen², Charlie Turner², Sophie Logan⁹, Abbie Harrison ², The COVID-19 Genomics UK (COG-UK) Consortium*, Sharon J. Peacock ^{1,2,3,4}, Meera Chand^{2,105} & Ewan M. Harrison^{1,2,4,10,105}✉

Mitigation of SARS-CoV-2 transmission from international travel is a priority. We evaluated the effectiveness of travellers being required to quarantine for 14-days on return to England in Summer 2020. We identified 4,207 travel-related SARS-CoV-2 cases and their contacts, and identified 827 associated SARS-CoV-2 genomes. Overall, quarantine was associated with a lower rate of contacts, and the impact of quarantine was greatest in the 16–20 age-group. 186 SARS-CoV-2 genomes were sufficiently unique to identify travel-related clusters. Fewer genomically-linked cases were observed for index cases who returned from countries with quarantine requirement compared to countries with no quarantine requirement. This difference was explained by fewer importation events per identified genome for these cases, as opposed to fewer onward contacts per case. Overall, our study demonstrates that a 14-day quarantine period reduces, but does not completely eliminate, the onward transmission of imported cases, mainly by dissuading travel to countries with a quarantine requirement.

¹University of Cambridge, Department of Medicine, Cambridge, UK. ²Public Health England, 61 Colindale Ave, London NW9 5EQ, UK. ³Cambridge University Hospital NHS Foundation Trust, Cambridge, UK. ⁴Wellcome Sanger Institute, Hinxton, Cambridge, UK. ⁵Quadram Institute Bioscience, Norwich Research Park, Norwich NR4 7UQ, UK. ⁶Imperial College London, Department of Infectious Disease Epidemiology, London, UK. ⁷University of Cambridge, Cambridge University Libraries, Cambridge, UK. ⁸Public Health England National Infections Service, Field Service, Leeds, UK. ⁹Public Health England, National Infections Service, Field Service, Nottingham, UK. ¹⁰University of Cambridge, Department of Public Health and Primary Care, Cambridge, UK. ¹⁰⁴These authors contributed equally: Andrew J. Page, Ulf Schaefer. ¹⁰⁵These authors jointly supervised this work: Meera Chand, Ewan M. Harrison. *A list of authors and their affiliations appears at the end of the paper. ✉email: dinesh.aggarwal@nhs.net; eh6@sanger.ac.uk

SARS-CoV-2 was first identified in Wuhan, China¹ in December 2019 and has since been imported into virtually every country and region in the world^{2,3}. Understanding and tracking the sources of importation between countries can give important information for policymakers, and for managing the pandemic, by informing policies aimed at reducing the further spread of virus⁴. It is particularly important now as countries aim to mitigate the introduction of highly transmissible variants of concern with potentially reduced vaccine efficacy^{5,6}. The available brakes on imported SARS-CoV-2 cases include travel bans, quarantine measures, and testing of returning travellers⁷. These can apply to all countries or targeted to high-risk countries, for variable durations, and with variable degrees of enforcement.

In England, from 17 March 2020 to 4 July 2020, the government advised against all non-essential travel worldwide⁸. Between 4 July 2020 and 1 February 2021, travel corridors to countries deemed to be low risk for COVID-19 disease (subject to assessment and change) were established in which returning travellers were no longer required to quarantine for 14 days (at home). Persons returning from countries outside this list (except for exemptions e.g. specific employment) were required to quarantine at home (Fig. 1). This policy aimed to reduce the impact of travel-related SARS-CoV-2 cases in England⁹ possibly through limiting onwards transmission of SARS-CoV-2¹⁰ and deterring travel to those countries. Upon identification of an imported case, contact tracing and quarantine/self-isolation measures can limit onwards transmission¹¹. The PHE Isolation Assurance Service identified up to 97% self-reported compliance with travel-specific quarantine¹². These data do not include countries exempt from quarantine, contact-tracing data or link to genomic data to evaluate travel-related clusters.

Studies from numerous countries have used genome sequencing to complement epidemiological investigations in order to characterise importations of SARS-CoV-2 (Supplementary Table 1). Primarily these are in-depth case reports on small datasets but demonstrate the utility of genomics combined with contact tracing^{13–35}. Here, we combine contact-tracing data from National Health Service (NHS) Test and Trace (T&T) for probable importation cases with genomes made available through the COVID-19 Genomics UK (COG-UK) consortium³⁶ to characterise the known imported cases of SARS-CoV-2 into England and the effectiveness of 14-day quarantine on onwards transmission.

In this work, we compare the number of contacts reported per case prior to diagnosis between individuals returning from a country with a requirement to quarantine after travel to those who did not need to quarantine on return. We then identify unique genomes from imported cases and associated clusters of infections in the COG-UK genomic surveillance dataset in the four weeks following index case identification. This onward transmission is compared when the index case returned from a country with a requirement to quarantine on return compared to countries without a requirement. Finally, we use the epidemiological data to investigate the origin of a divergent cluster of SARS-CoV-2 cases identified using genomics.

Results

Cases identified. Between 27 May 2020 and 13 September 2020, using contact-tracing data for all individuals who had tested positive for SARS-CoV-2 between those dates, we identified 4207 international travel-related cases in England. These individuals reported a total of 18,856 contacts. During this period, we identified 105,794 non-travel related cases that reported 233,182 contacts.

From the travel-related cases, 888 sequenced genomes were available for comparison to all UK genomic data (see Fig. 1 and Methods for details of case definition and identification,

Supplementary Tables 2 and 3 show the case characteristics). Sequencing of community and hospital cases across the UK was carried out with the aim of providing approximately equal geographical coverage, as much as possible.

Return from European countries accounted for 85.9% (3612/4207) of travel-related cases; 51.2% (2155/4207) had visited one of Greece (21.0%, 882/4207), Croatia (16.3%, 685/4207) or Spain (14.0%, 589/4207) (Fig. 1 and Supplementary Table 4). For 284 cases the country of travel was unclear or unknown. Travel restrictions were first eased on 04/07/2020; only 2.9% of travel-related cases identified in this study were recorded before this date. For the countries associated with the highest numbers of imports, the number of cases per day imported from each country along with the timing of travel restriction to that country is shown in Fig. 2. Geographical variations in imported cases across England were apparent, with the greatest number (28.6%, 1205/4207) in Greater London (representing ~15% of the population of England) (Fig. 1 and Supplementary Table 3).

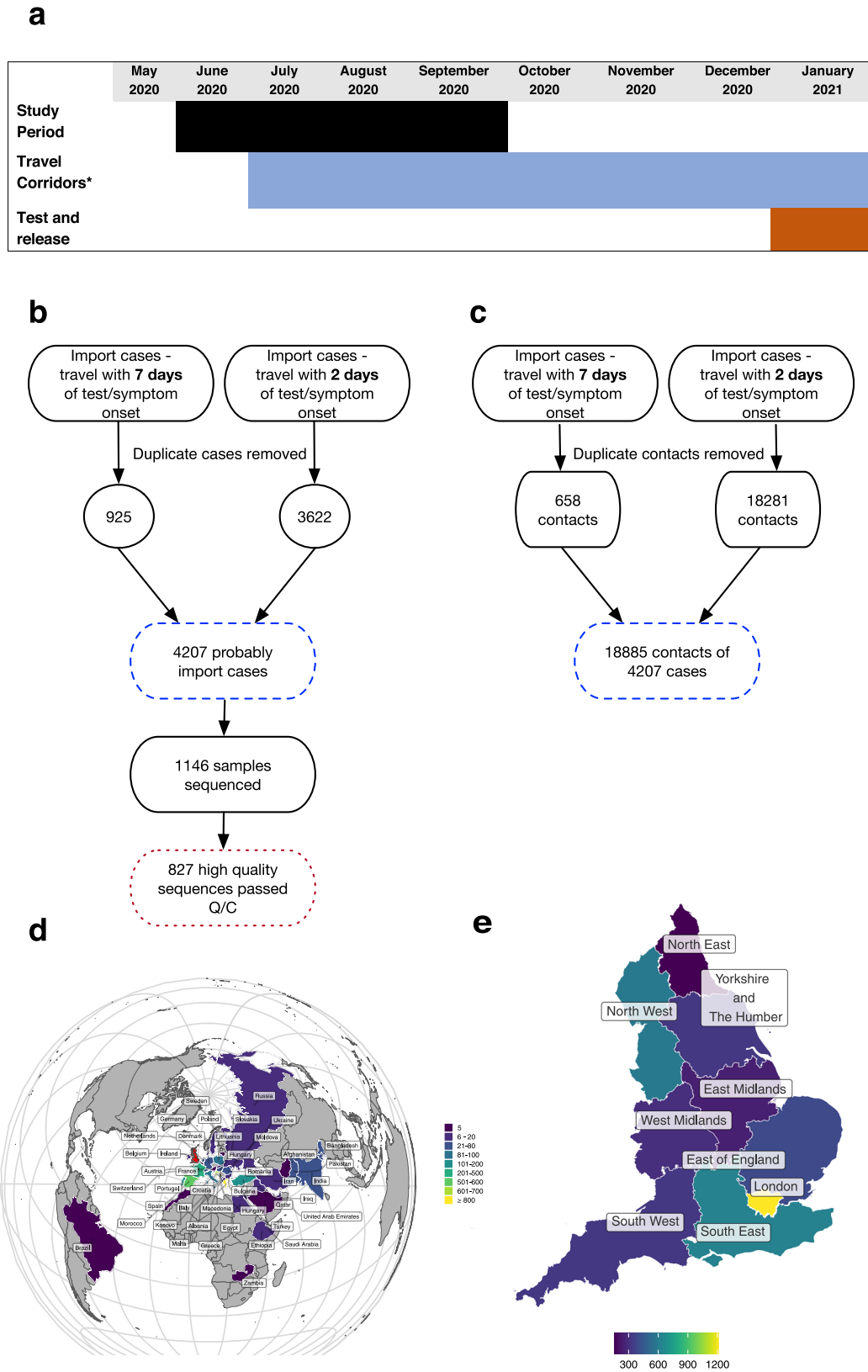
Contacts per case. The median number of reported contacts per travel-associated case was 3 (IQR 1–5), with 22% reporting no contacts, while some individuals reported a very large number of contacts, 9% reported more than 10, 3% more than 20, 0.4% more than 50, with a maximum of 172.

Of the imported cases, 2010 were imported from a country with a quarantine requirement at the time of return, whereas 1900 were not required to quarantine on return. For 297 cases quarantine status could not be determined. The number of contacts was higher for cases without a travel restriction (mean = 6.0, median = 3, IQR = 1–7) compared to cases with a travel restriction (mean = 3.0, median = 2, IQR = 1–4).

Using a negative binomial regression model, after adjusting for potential confounding factors of age, sex, date of the test, destination, and ethnicity, travelling from a country requiring quarantine on return was associated with an estimated reduction in the number of contacts of 40% (rate ratio (R.R.) = 0.60, 95% CI = 0.37–0.95; $p = 0.03$). Statistical modelling is fully described in the “Methods” section. Using this model, the estimated marginal mean number of reported contacts (adjusting and averaging over all covariates; age, sex, date of the test, destination, and ethnicity) was 5.85 (95% CI = 3.7–9.3) when no quarantine was required compared to 3.50 (95% CI = 3.0–4.0) when travellers were required to quarantine. To address possible bias from a small number of cases with a large number of contacts, we recorded (top-coded) all cases with more than 10 contacts as corresponding to 10; the estimated rate ratio was slightly attenuated to 0.68 (0.48–0.98; $p = 0.036$) for the number of contacts per case for individuals travelling from a country requiring quarantine on return compared to those not requiring quarantine.

The number of contacts per case varied significantly with age group and over time (Fig. 3 and Supplementary Table 6). The number of contacts per case was greatest in the 16–20 age group who travelled to countries with no requirement for quarantine, with a marginal mean of 9.0 (95% CI = 5.6–14.5) but reduced to 4.7 (95% CI = 3.9–5.7) when quarantine was required—similar to other age-groups.

After adjusting for all other covariates the reported numbers of contacts per imported case was lower in September compared to May, June and July, whether or not a requirement to quarantine was in place. Following this observation of reduced contacts over time among travel-related cases, we compared this to the number of reported contacts over time among the remaining population who did not travel to ensure this was not a general trend due to other COVID-19 measures. Among 105,794 cases recorded in our study period that were not associated with travel, 28,564 (27%) were excluded due to poor



data quality. Among the remaining 77,230 cases we did not find a corresponding decrease over time in the number of contacts, with a mean of 1.6 contacts per case in May/June 2020 and around 2.3 contacts per case for the remaining period.

Onward transmission and genomic analysis. We next sought to quantify onward transmission from an imported case using genomics. High-quality sequencing data were available for 827/4207 (19.7%) cases (Fig. 1) and demographics of the sequenced

Fig. 1 Case ascertainment and distribution during the study period. **a** Timeline of the study period (27 May 2020 to 13 September 2020) and associated policy changes on travel introduced in England. Travel-related quarantine measures were assigned on a country by country basis from 4 July 2020. Travellers returning from countries that were on the ‘closed travel corridors list’⁵⁸ were required to quarantine for 14 days (‘reduced to 10 days on 15/12/20), or from the 15th December 2020, choose to self-isolate for 5 days and then pay for a SARS-CoV-2 diagnostic test (test and release). **b** Flow diagram and map of travel-related cases ascertained from Test and Trace data and subsequent genome availability. Cases were defined as ‘highly probable’ and ‘probable’. ‘Highly probable’ travel-related cases were defined as individuals who reported international travel as an activity in the two days before symptom onset/testing. On 12/08/2020 the additional facility to report international travel in the 7 days prior to symptom onset/testing became available, and also included in this study and defined as ‘probable’ travel-related cases. **c** Flow diagram relating contacts ascertained of cases from Test and Trace data. **d** Countries where importations originated. Countries with less than five importations were excluded for confidentiality reasons. **e** Destinations of imported cases within England. Areas with less than three cases have been excluded. Q/C quality control.

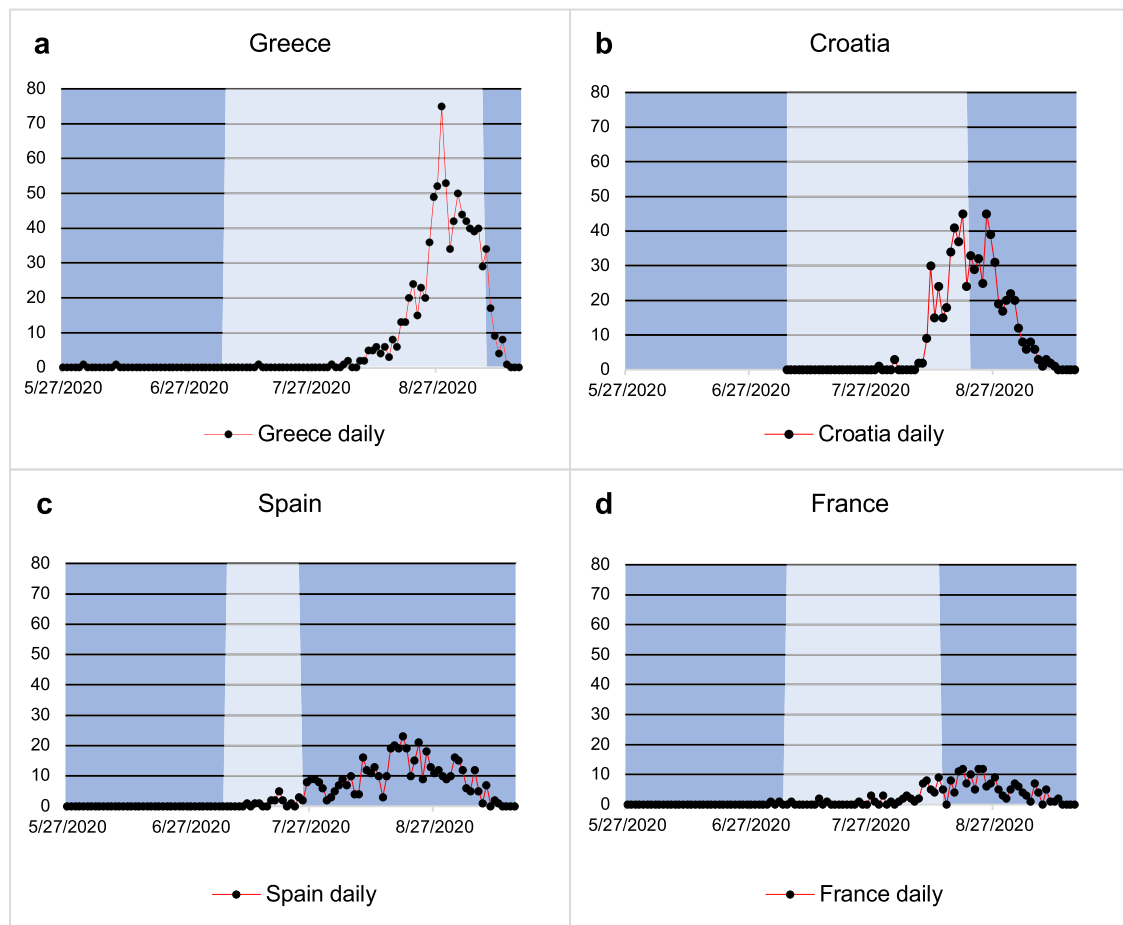


Fig. 2 Frequency of importations overtime for the top 4 most common countries of travel reported by individuals testing positive for SARS-CoV-2 during the study period. **a–d** SARS-CoV-2 case numbers in returning travellers by the four most popular countries of travel reported by cases representing 2379/4207 (56.5%) of known travel-related cases. The light-shaded areas represent the period of time when the countries had an open ‘travel-corridor’ so did not have mandatory 14-day quarantine on return in place.

cases were broadly similar to the entire travel-related cohort (Supplementary Table 3).

Important genomes and onward transmission. To monitor onward spread we identified 186/827 (22.4%) imported cases with SARS-CoV-2 genomes that were sufficiently unique, as defined by their status as extinct or genetically distinct within a sub-lineage (see Methods). Of these, 146/186 isolates had not been sampled in the entire UK dataset in the 4 weeks prior, while a further 40 isolates were more than 3 single nucleotide polymorphisms (SNPs) to their closest matching sequence in the existing UK dataset, both suggesting genuinely new importations of this genotype.

Using an SNP-matrix of imported genomes and associated travel-related metadata, we defined the number of importation events per imported genotype; these ranged from 1 to 39. The majority (119/186; 64%) of genomes were identified only once in imported cases, with 33 (18%) identified twice, 22 (12%) between 3 and 10 times and the remaining 12 (6%) between 11 and the maximum of the observed number of 39.

To compare the effect of the requirement to quarantine on the subsequent spread of likely imported cases, the entire COG-UK dataset was interrogated to identify isolates within 2 SNPs of these distinct imported cases identified up to 4 weeks after the index importation case. There was variation in the number of subsequent (up to 4 weeks later) cases matching each genome

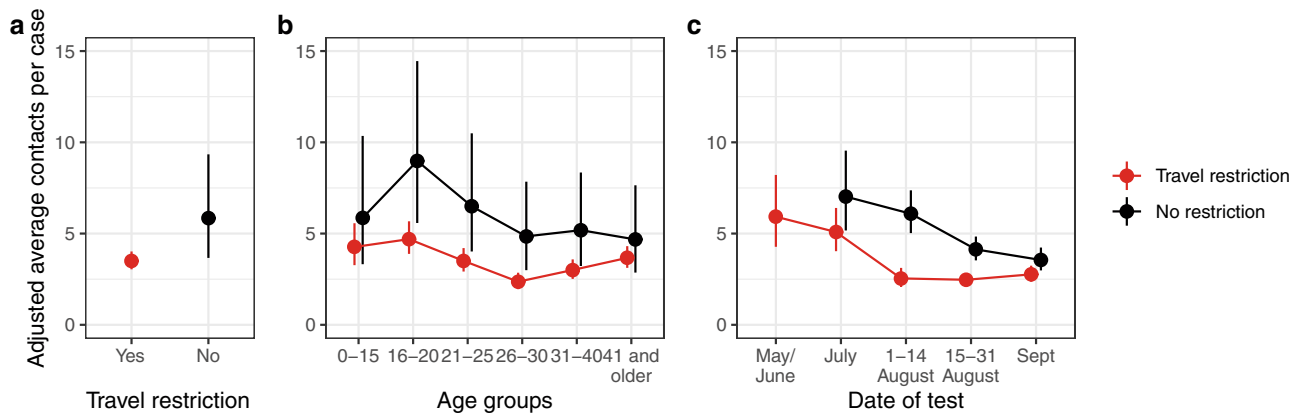


Fig. 3 The effect of travel restriction (14-day quarantine) on contacts per imported case of SARS-CoV-2. The estimated marginal mean number of contacts per imported case **a** overall, **b** by age-group and **c** by date of test comparing countries with travel restriction guidance (closed ‘travel-corridors’) in place and those without (open ‘travel-corridors’). All points are estimated marginal means and are provided with 95% confidence intervals.

(median 0; range 0–210, IQR 0–1). The majority of genomes 125/186 = 67% were not linked to any subsequent cases, 17 (9.1%) and 8 (4.3%) were linked to one or two cases, with the remaining 36 (19%) being matched to larger numbers of subsequent cases, with a small number of imported cases corresponding to large numbers of subsequent cases, including 6 that matched to at least 50 later cases.

Association between travel restriction and onward transmission.

To explore the association between onward transmission and travel restriction we first excluded cases returning before 14 July 2021 to ensure the time periods between index cases with and without a travel restriction overlapped. The proportions of imported genomes matching any subsequent detected case and the number of new cases where at least one is detected in this group are shown in Fig. 4. Overall, 56/168 (33%) of genomes from cases that were genetically unique were detected in at least one subsequent case within the subsequent four weeks. Among genomes identified from a country where quarantine requirement was in place, 25% of (20/81) were detected in at least one subsequent case, compared to 41% (29/71) when cases were imported from a country without a requirement to quarantine (Fig. 4a). The destination country for 16 index cases was unknown.

The number of subsequent cases detected during the 4 weeks since the unique index case increased from a mean of 1.2 new cases when quarantine was required to 11.3 cases where there was no requirement, mainly driven by the fact that all of the nine genomes which went on to match more than 20 subsequent cases had an index case returning from a country without travel restriction (Fig. 4b). However this difference can be explained entirely by the number of importations for each genome; these genomes all had high numbers of independent importations, and genomes with a high number of importations always had an index case returning from countries with no travel restriction in place at the time (Fig. 4c).

To test the statistical significance of observed effects, a series of negative binomial regression models were fitted. In the four weeks following the index case, fewer genomically linked cases were reported when the index case was imported from a country with a requirement to quarantine compared to cases from a country with no requirements (unadjusted R.R. = 0.11, 95% CI = 0.04–0.28), but this effect was entirely explained when the number of imported cases for each genome is included as an ‘offset’ in the model and adjusting for the date (R.R. = 0.83, 95% CI = 0.35–1.92; $p = 0.655$).

There was some evidence that imported cases with higher numbers of contacts for the index case gave rise to more cases in the subsequent month, however, the number of contacts was only known for the index case. This effect is also explained by the number of importations; there is a positive correlation (Spearman’s $\rho = 0.18$, $p = 0.018$) between the number of contacts reported by the index case and the number of independent importations of each genome (Supplementary Fig. 1), and in regression models, once we adjust for the number of importations there is no evidence for an association between a number of contacts of the index case and subsequent linked cases. However, it is possible that the number of contacts of other imported cases in each group remains an important factor.

Genomic identification of a large imported cluster. In order to demonstrate the utility of genomics in identifying a probable travel-related cluster of SARS-CoV-2 cases in the England de novo, we ran the Polecat Clustering tool (<https://cog-uk.github.io/polecat>) on 14 September 2020 (including SARS-CoV-2 cases in the COG-UK dataset up to this date). An outlier cluster was observed (Supplementary Fig. 4). This cluster (UK1897) was associated with high diversity with a long stem length compared to samples from the UK, suggesting that this lineage evolved outside the UK. The geographic distribution of this lineage is demonstrated in Supplementary Fig. 5, likely representing multiple importations into the UK. This cluster contained the D614G mutation but no others associated with increased transmission. The root of the cluster was associated with a Swiss phylotype when linked to data in GISAID. During the course of the study period (4 August 2020 to 14 September 2020), there were 304 genomes corresponding to this cluster. These could be linked to 238 individuals, of whom 159 could be linked to a contact-tracing record. Out of 159, 143 had contact-tracing information indicating international travel or not. Out of 143, 72 (50.3%) individuals had recently returned from abroad and were associated with, 10 dispersed European countries (4 individuals had traveled to more than 1 European Country) and most commonly Croatia (35/72, 48.6%) (Supplementary Fig. 6). A further four cases were identified as contacts of individuals who had reported travel to mainland Europe. There is a trend towards an increased proportion of cases that do not report travel over time, and possibly representing dispersion and onwards transmission locally of this lineage (Supplementary Fig. 7).

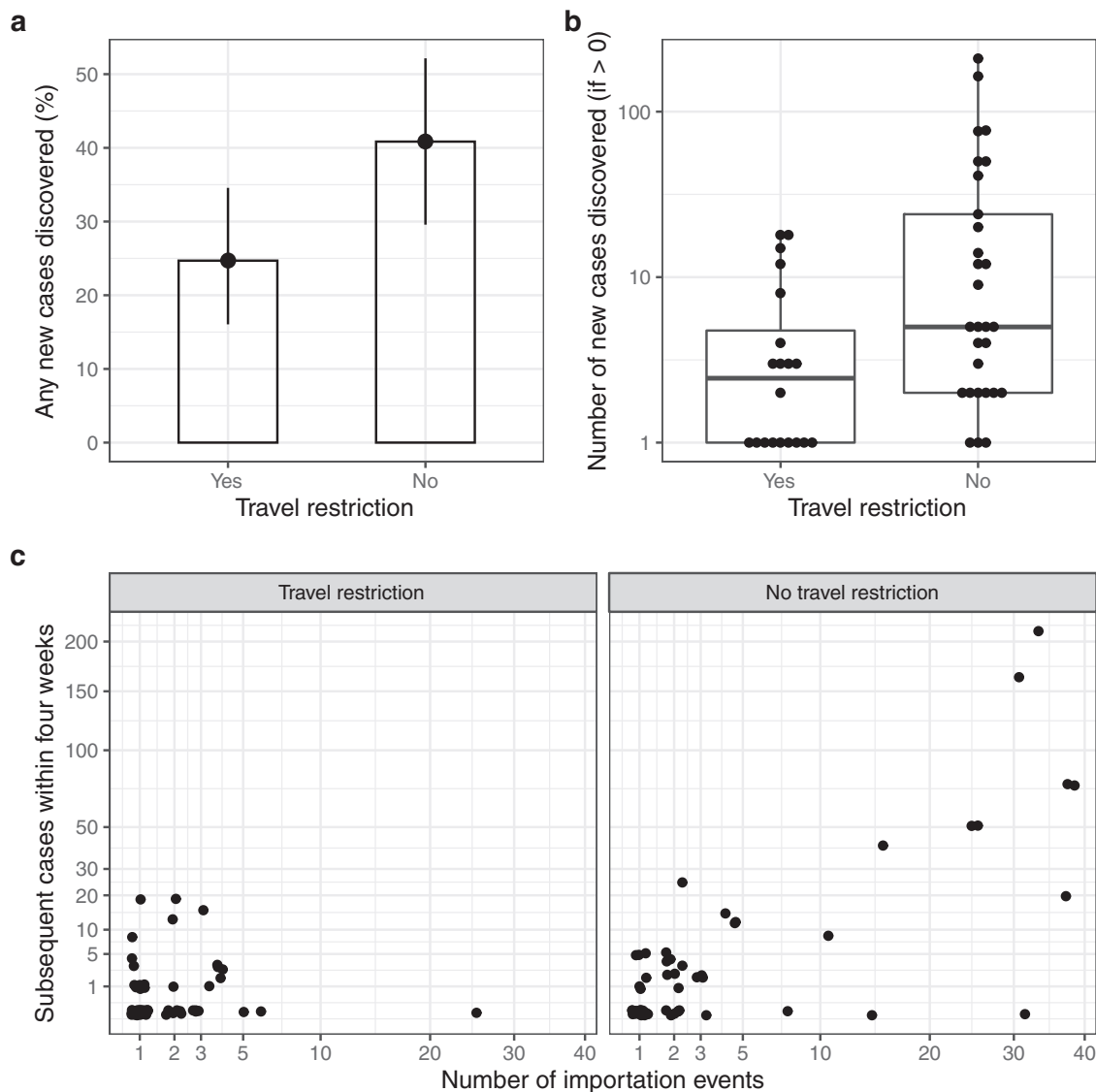


Fig. 4 The effect of travel restrictions (14-day quarantine) on the subsequent spread of likely imported cases as determined by genomics. **a** The proportion of imported cases with any matching genome detected over the four weeks following index test result. Error bars correspond to bootstrapped confidence intervals, $n = 81$ and 71 . **b** The number of genomes matching the index case, with zeros excluded. The midline of the boxplot represents the median value; the lower limit of the box represents the first quartile (25th percentile), and the upper limit of the box represents the third quartile (75th percentile); the whiskers (upper and lower) extend to the largest and smallest value from the box, no further than $1.5 \times \text{IQR}$ from the box. **a, b** Compare countries with travel restriction guidance (closed ‘travel-corridors’) in place and those without (open ‘travel-corridors’). **c** The number of cases and number of importation events for each imported genome, stratified by whether the index case had returned from a country with a travel restriction in place.

Characteristics of imported genomes. The 827 imported genomes reflected 238 UK lineages (see Supplementary Materials), of which 214 were seen fewer than 5 times (142 singletons) and 24 were seen 5 or more times (Supplementary Table 8). The most commonly observed were UK5 (152 genomes, 18.4%) and UK1897 (73 genomes, 8.8%). There were 39 global lineages within the genomes. The most commonly observed lineages were B.1.1 (159 genomes, 19.2%) and B.1.177 (128 genomes, 15.5%) (Supplementary Tables 9 and 10). Potentially functionally important mutations were also identified (Supplementary Table 11 and Supplementary Fig. 8): D614G, 824/827 (99.6%) cases; N439K, 65/827 (7.86%) of cases; A222V, 131/827 (15.84%) of cases. $\Delta\text{H69/V70}$ was identified in 53 cases associated with lineage B.1.258. We evaluated the introduction of A222V (B.1.177) over time, demonstrating a clear epidemiological link to Spain through contact tracing (Supplementary Fig. 10). By the end of the study period, this variant was introduced from

16 separate countries indicating dispersion across Europe (Supplementary Fig. 11) corroborating findings by Hodcroft et al.³⁷. The mutations co-occur, with the proportion of cases represented by these combinations varying over time (Supplementary Fig. 12).

Discussion

Here, we provide evidence, through the analysis of both contact-tracing data and the use of genomics, that a mandatory 14-day quarantine was associated with fewer contacts for returning travellers with SARS-COV-2, and less onward transmission of imported cases, with the reduced transmission, likely mediated through fewer individual importations of each genome. From 27 May 2020 to 13 September 2020, 85.9% of importations of SARS-COV-2 into England were from European countries with three countries, Greece, Croatia, and Spain, accounting for 51.2% of all imported cases. Along with the requirement to quarantine or not,

age was a significant determinant of onwards contacts, with younger age groups reporting more contacts, but the effect of a requirement to quarantine on the number of contacts was observed across all age groups. We have shown that after a period of national lockdown, systematic monitoring of imported genomes can identify sequences that are sufficiently unique and provide utility for monitoring onwards transmission.

Whilst the study period covers nearly 5 months, the importations were concentrated after the implementation of travel corridors; prior to this date travel was not advised³⁸. The peaks for imports for each country occur at different times and with different epidemic curves, likely affected by both patterns of travel as well as the prevalence of disease and local regulations within each country. For the most common destinations, barring Spain, imported cases appear to reduce after the closing of a travel corridor and subsequent requirement to quarantine. The majority of importations from Greece came at the end of August and continued into September; there was no requirement to quarantine for travellers returning from Greece during this time period and it was the source of the greatest imported SARS-CoV-2 cases during the study period. This highlights the need for active surveillance of imported cases of SARS-CoV-2 for the introduction of requirements to quarantine in a timely manner. London accounts for 15.4% of the population in England³⁹ and observed 11.4% (12011/105794) SARS-CoV-2 infections during the study period; the region however accounted for 28.6% of imported SARS-CoV-2 cases, possibly reflecting a younger and more diverse demographic with cultural/family links abroad, and with a concentration of international businesses and airports. The reported effective reproduction number (R_e) in London had a minimum lower-bound value of 0.6 and an upper-bound value of 1.3 during the study period which was comparably similar to the respective values of 0.7 and 1.2 observed for England⁴⁰. This potentially indicates imports are unlikely to have had a substantial impact on onward infection rates in this region.

The number of onwards contacts were significantly lower when the traveller was required to quarantine highlighting the effectiveness of this policy. Age was also a significant determinant of onwards contacts, with the 16–20-year-old age group representing the greatest number of travel-related cases and the greatest number of onwards contacts per case. This identifies an opportunity to direct public health awareness campaigns to younger travellers, with the intention to promote behaviours that will reduce the risk of SARS-CoV-2 acquisition and enhance compliance with a quarantine on return.

We observed a reduction in the number of contacts per case over the Summer 2020 period that appeared specific to the travel-related SARS-CoV-2 cases as it was not replicated in the non-travel related cases. There was also no apparent change in successful contact tracing to explain this difference (Supplementary Fig. 13). We speculate this observation may be related to a change in traveller behaviour (e.g. due to rising cases in the England or destination countries) or changes in types of traveller (e.g. travellers visiting family versus those for occupational reasons or those with dependents versus those without). Understanding these temporal changes in traveller behaviour requires formal investigation with more detailed epidemiological information available and replication in other countries. More broadly, we do find non-travellers reported fewer contacts per case when compared to travellers which may reflect a more sociable cohort, and therefore one that may benefit from targeted public health messaging to reduce transmission risk.

The use of genomic sequencing allowed the identification of a cohort of unique genomes that could be monitored for cluster growth. The cluster size for genomes that were imported from a country without a requirement to quarantine on return was

significantly higher than those related to countries with mandatory quarantine in place providing further assurance on the effectiveness of quarantine policy on reducing travel-related SARS-CoV-2 cases. This finding was explained by several large clusters, all of which came from countries with no quarantine requirement at the time and with high numbers of individual importation events. With the number of importation events per cluster taken into account, we did not observe an effect of quarantine on subsequent cluster size suggesting the largest effect of travel-related quarantine is through a net reduction in travel-related importations.

The Polecat Clustering Tool (<https://cog-uk.github.io/polecat>) highlighted a large cluster that developed largely through travel to Croatia. This analysis shows that programmatic analysis of genomics data can identify putative importation clusters. Integration with contact-tracing information was vital for the true picture of the sources of introduction and the subsequent spread, due to the bias of SARS-CoV-2 sequencing globally⁴¹. In this instance, an introduced lineage was associated with widespread dispersal and onward transmission during a period when England had limited social distancing measures^{42,43}. The lineage, B.1.160, associated with this cluster is not associated with increased transmissibility but this method for the detection of expanding imported clusters could be useful for the investigation of newly introduced variants of concern.

Our study has several important limitations. The COG-UK dataset has a limited sequencing coverage across England meaning cluster sizes detected will under-estimate absolute numbers and there is a possibility of unsampled transmission chains despite our use of four SNPs to identify divergent clusters (Supplementary Fig. 14). The earliest reliable data available to identify if individuals were required to quarantine on returning to England from travel abroad was the date of case sampling. Our study evaluates a period of time following a national lockdown with highly restricted movement across borders which likely exaggerated the diversity of imported genomes compared to lineages circulating in England. Additionally, the quarantine guidance at this time was of 14 days; shorter periods may be as efficacious and/or when combined with testing⁴⁴. Outcomes such as travel and the number of contacts are self-reported; reporting bias is mitigated through mandatory completion of a passenger locator form to assist identification of returning travellers, while travel-related cases are seen as higher risk and therefore referred to local public health agencies for targeted contact tracing. For genomic analysis, only the destination country and number of contacts for the index case in each cluster are used, irrespective of the number of imports. Finally, there will be an artificial reduction in cases at the end of the study period when accounting for the case incubation period, testing and report, with data provided 3 days after study close.

Overall, we present an integrated epidemiological and genomic evaluation of the largest dataset of confirmed SARS-CoV-2 imported cases into England (or any other country) to our knowledge. This study provides evidence for the effectiveness of 14-day quarantine in reducing contacts, and reducing, but not entirely preventing onward transmission of imported cases, through reducing the number of importations. Our data highlights the possibility of targeted public health campaigns to reduce SARS-CoV-2 importations and onwards transmission. In conclusion, this study demonstrates how routine genomic epidemiology of travel-related cases could be used to monitor SARS-CoV-2 import cases to enable rapid refinement of travel policies.

Methods

Ethics. The COG-UK study protocol was approved by the Public Health England Research Ethics Governance Group (reference: R&D NR0195). Public Health

England affiliated authors had access to identifiable Cambridgeshire community case data. This data was processed under Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002—permitting the processing of confidential patient information for communicable disease and other risks to public health and as such, individual patient consent is not required. Other authors only had access to anonymised or summarised data.

Contact tracing and case identification. Contact-tracing data was obtained from Test and Trace (T&T). All cases and contacts had a field for demographic data, but this was not always reported (Supplementary Tables 3 and 4). ‘Highly probable’ travel-related cases were defined as individuals who reported international travel as an activity in the two days before symptom onset/testing. On 12/08/2020 the additional facility to report international travel in the seven days prior to symptom onset/testing became available, and also included in this study and defined as ‘probable’ travel-related cases.

Cases were asked to provide details of all contacts for activities in the 2 days prior to symptom-onset/testing (whichever is earliest) up to the time of completing the system in which contacts were gathered. Though there can be a discrepancy in the time taken for individuals to complete the contact tracing system, this is not expected to result in a material change in contact numbers as they are expected to be self-isolating after symptom-onset or testing positive. If any contacts become cases they would then also be included in T&T data as a case separately, but if they did not report direct travel themselves, then they would not meet the definition for a travel-associated case.

A contact is defined as an individual who a case has had face-to-face contact with (less than 1 metre away), spent more than 15 min within 2 m of, travelling in a car or other small vehicle with, or sat close to them on a plane

Case identification from T&T data. Data included free-text destination city or country. A free text country and city search with a custom python script on travel-related T&T were used to identify destination country. Results and remaining entries were manually checked and corrected (see Supplementary materials for more details).

Requirement for quarantine. Persons returning from countries where travel-related quarantine was not mandated (except for exemptions e.g. specific employment) were required to quarantine for 14 days at home. The package of measures used to help enforce this requirement included the need to complete a Passenger Locator Form prior to arrival in England, spot checks by the PHE Isolation Assurance Service, and referral to the police through the Border Force Criminal Justice Unit who may issue fixed penalty notices.

Clinical samples, genome sequencing and quality control. Clinical samples were collected passively as part of national SARS-CoV-2 testing. This included both community testing through lighthouse labs (satellite SARS-CoV-2 testing laboratories) and testing through hospital diagnostic labs. Samples were sequenced at one of seventeen COG-UK sequencing sites (Fig. 1). The samples were prepared for sequencing using either the ARTIC⁴⁵ or veSeq⁴⁶ protocols and were sequenced using Illumina or Oxford Nanopore platforms. All samples were uploaded to and processed through COVID-CLIMB pipelines^{47,48}. Genomes were aligned to the Wuhan Hu-1 reference genome (Genbank accession code: MN908947.3). Genomes that contained more than 10% missing data were excluded from further analysis to ensure high-quality phylogenetic analysis.

Lineages and minor variants. Global and UK Lineages⁴⁹ were assigned to each genome using Pangolin (<https://github.com/cov-lineages/pangolin>) with analysis performed on COVID-CLIMB⁴⁸. Minor variants were pre-defined within the COG-UK database using type_variants (https://github.com/cov-ert/type_variants).

Identification of extinct and unique genomes. The 827 high-quality travel-related genomes were compared to the COG-UK dataset on 16/10/2020. Genomes were only compared to other genomes with the same UK lineage assigned by COG-UK since we assume that no relatedness relevant to transmission exists between genomes of different UK lineages. A unique genome in the community was deemed to be one that was known to be from a travel-related case and that was either: a UK lineage that had not been sampled in the previous 4 weeks in the UK or was more than 3 SNPs distance to the closest relative in the COG-UK dataset.

Within the same UK lineage, we then identified those genomes sampled within 4 weeks prior to the genome of interest. We determined the SNP distance between the sequence of interest and these genomes. Unique genomes were compared to sequences that were generated in the COG-UK dataset within 2 and 4 weeks after their sampling date, to identify samples with the same UK lineage and within 2 SNPs. These would represent onward transmission or further introductions of similar genomes. The analysis was run with an in-house custom Python script developed by US and RM. Further detail in Supplementary materials.

Identification of multiple introductions of a unique genome. We combined the available travel-related epidemiological data and genomic data to identify the

number of importations representing the travel-related clusters generated above. We used SNP-dist (version 0.7.0) (<https://github.com/tseemann/snp-dists>) to identify the SNP distances between an alignment 827 high-quality imported genomes. These genomes were aligned with MAFFT (version 7.471)⁵⁰, outside of CLIMB-COVID pipelines, with minor differences in SNP differences to the entire COG-UK alignment expected. We then identified imported genomes that were within 2 SNPs of the 186 unique imported genomes in the 4 weeks subsequent to the unique imported case being sampled. This represented the number of importations of that genome in the 4-week period of interest corresponding to each unique travel-related cluster identified in the analysis above.

Identification of a travel-related SARS-CoV-2 cluster. We used the Polecat clustering tool (<https://cog-uk.github.io/polecat>) to systematically identify outliers in COG-UK genomic dataset and link to contact-tracing data.

Statistical analysis. Statistical models for the number of contacts per case and the number of onward cases per imported genotype were estimated using the glmmTMB package (version 1.0.1)⁵¹ with marginal means and effects calculated using the emmeans package (1.5.2-1)⁵² for R (version 4.0.2)⁵³. Figures were generated using R (version 4.0.2) and Microsoft Excel (version 1908).

The number of contacts per case was modelled using negative binomial regression, to estimate the effect of travel-related quarantine, and whether this varied by age group, sex of the index case and calendar date. Travel destination and ethnic group were included as random effects. Negative binomial regression models were also used for the number of onward cases per imported genome, using calendar date as a covariate and the natural log of the number of importation events for each genome as an offset variable. Model validation was performed by simulation from estimated models and comparing the distribution of observed and modelled outcomes. For the contacts modelling the initial negative binomial regression model did not reflect the number of large outliers seen in the original data, therefore additional models were estimated using data with these observations dropped or top-coded to check that the reported estimates of effects were not driven by these observations. For the genomic data, the simulated and observed distributions were closely aligned.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Assembled/consensus genomes are available from GISAID⁵⁴ subject to minimum quality control criteria. Raw reads are available from European Nucleotide Archive (ENA)⁵⁵ under accession PRJEB37886. ENA accession codes for travel-related genomes used in this study are available in supplementary materials (Supplementary Data 1) and available from GitHub at <https://github.com/COG-UK/travel-quarantine>⁵⁶. The Genbank accession code for the reference genome uses is MN908947.3. All genomes, phylogenetic trees, basic metadata are available from the COG-UK consortium website (<https://www.cogconsortium.uk/data>). Extensive aggregated metadata has been made available in supplementary files. Genomes accessed through GISAID used in this study are provided in the Supplementary Information file entitled ‘GISAID acknowledgement table’. For confidentiality reasons⁵⁷, extended metadata are under restricted access; requests for access should be directed to corresponding authors and specifically for Public Health England data, to the Public Health England office of data release (<https://www.gov.uk/government/publications/accessing-public-health-england-data/about-the-phe-odr-and-accessing-data>) with an estimated 60 working days turnaround time.

Code availability

Custom code used in this analysis is available at <https://github.com/COG-UK/travel-quarantine>⁵⁶. Please direct further queries to the corresponding authors.

Received: 24 May 2021; Accepted: 18 January 2022;

Published online: 23 February 2022

References

- Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
- Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 564–570 (2020).
- Wells, C. R. et al. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proc. Natl. Acad. Sci. USA* **117**, 7504–7509 (2020).
- da Silva Filipe, A. et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. *Nat. Microbiol.* **6**, 112–122 (2021).

5. Abdool Karim, S. S. & de Oliveira, T. New SARS-CoV-2 variants—clinical, public health, and vaccine implications. *N. Engl. J. Med.* **384**, 1866–1868 (2021).
6. Gupta, R. K. Will SARS-CoV-2 variants of concern affect the promise of vaccines? *Nat. Rev. Immunol.* <https://doi.org/10.1038/s41577-021-00556-5> (2021).
7. Johansson, M. A. et al. Reducing travel-related SARS-CoV-2 transmission with layered mitigation measures: symptom monitoring, quarantine, and testing. *BMC Med.* **19**, 94 (2021).
8. Office, F. C. *Travel Advice: Coronavirus (COVID-19)*. <https://www.gov.uk/guidance/travel-advice-novel-coronavirus> (2020).
9. Transport, D. f. *Travel Corridors*. <https://questions-statements.parliament.uk/written-statements/detail/2020-07-06/HCWS338> (2020).
10. Chiew, C. J., Li, Z. & Lee, V. J. Reducing onward spread of COVID-19 from imported cases: quarantine and ‘stay at home’ measures for travellers and returning residents to Singapore. *J. Travel. Med.* <https://doi.org/10.1093/jtm/taaa049> (2020).
11. Kucharski, A. J. et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1151–1160 (2020).
12. Force, B. *Data on Health Measures at the UK Border* <https://www.gov.uk/government/publications/data-on-health-measures-at-the-uk-border/data-on-health-measures-at-the-uk-border> (2020).
13. Au, C. H. et al. Genome Sequences of SARS-CoV-2 Strains Detected in Hong Kong. *Microbiol. Resour. Announc.* <https://doi.org/10.1128/MRA.00697-20> (2020).
14. Bohmer, M. M. et al. Investigation of a COVID-19 outbreak in Germany resulting from a single travel-associated primary case: a case series. *Lancet Infect. Dis.* **20**, 920–928 (2020).
15. Cohen-Gihon, I. et al. Coding-complete genome sequences of two SARS-CoV-2 isolates from early manifestations of COVID-19 in Israel. *Microbiol. Resour. Announc.* <https://doi.org/10.1128/MRA.00677-20> (2020).
16. Jesus, J. G. et al. Importation and early local transmission of COVID-19 in Brazil, 2020. *Rev. Inst. Med. Trop. Sao Paulo* **62**, e30 (2020).
17. Nascimento, V. A. D. et al. Genomic and phylogenetic characterisation of an imported case of SARS-CoV-2 in Amazonas State, Brazil. *Mem. Inst. Oswaldo Cruz* **115**, e200310 (2020).
18. Du, P. et al. Genomic surveillance of COVID-19 cases in Beijing. *Nat. Commun.* **11**, 5503 (2020).
19. Garces-Ayala, F. et al. Full genome sequence of the first SARS-CoV-2 detected in Mexico. *Arch. Virol.* **165**, 2095–2098 (2020).
20. Giandhari, J. et al. Early transmission of SARS-CoV-2 in South Africa: An epidemiological and phylogenetic report. *Int J Infect Dis* **103**, 234–241 (2021).
21. Giovanetti, M., Benvenuto, D., Angeletti, S. & Ciccozzi, M. The first two cases of 2019-nCoV in Italy: where they come from? *J. Med. Virol.* **92**, 518–521 (2020).
22. Gomez-Carballa, A. et al. Phylogeography of SARS-CoV-2 pandemic in Spain: a story of multiple introductions, micro-geographic stratification, founder effects, and super-spreaders. *Zool. Res.* **41**, 605–620 (2020).
23. Gong, Y. N. et al. SARS-CoV-2 genomic surveillance in Taiwan revealed novel ORF8-deletion mutant and clade possibly associated with infections in Middle East. *Emerg. Microbes Infect.* **9**, 1457–1466 (2020).
24. Jia, Y. et al. Characterization of eight novel full-length genomes of SARS-CoV-2 among imported COVID-19 cases from abroad in Yunnan, China. *J. Infect.* **81**, e96–e98 (2020).
25. Kouriba, B. et al. First phylogenetic analysis of Malian SARS-CoV-2 sequences provides molecular insights into the genomic diversity of the Sahel Region. *Viruses* <https://doi.org/10.3390/v12111251> (2020).
26. Kumar, P. et al. Integrated genomic view of SARS-CoV-2 in India. *Wellcome Open Res.* **5**, 184 (2020).
27. Liu, J., Huang, J. & Xiang, D. Large SARS-CoV-2 outbreak caused by asymptomatic traveler, China. *Emerg. Infect. Dis.* <https://doi.org/10.3201/eid2609.201798> (2020).
28. Lu, J. et al. Genomic epidemiology of SARS-CoV-2 in Guangdong Province, China. *Cell* **181**, 997–1003 e1009 (2020).
29. Manning, J. E. et al. Rapid metagenomic characterization of a case of imported COVID-19 in Cambodia. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.03.02.968818> (2020).
30. Marquez, S. et al. Genome sequencing of the first SARS-CoV-2 reported from patients with COVID-19 in Ecuador. Preprint at *medRxiv* <https://doi.org/10.1101/2020.06.11.20128330> (2020).
31. Puenpa, J. et al. Molecular epidemiology of the first wave of severe acute respiratory syndrome coronavirus 2 infection in Thailand in 2020. *Sci. Rep.* **10**, 16602 (2020).
32. Seemann, T. et al. Tracking the COVID-19 pandemic in Australia using genomics. *Nat. Commun.* **11**, 4376 (2020).
33. Sekizuka, T. et al. SARS-CoV-2 genome analysis of Japanese travelers in Nile River Cruise. *Front. Microbiol.* **11**, 1316 (2020).
34. Stange, M. et al. SARS-CoV-2 outbreak in a tri-national urban area is dominated by a B.1 lineage variant linked to a mass gathering event. *PLoS Pathog.* **17**, e1009374 (2021).
35. Rockett, R. J. et al. Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nat. Med.* **26**, 1398–1404 (2020).
36. The COVID-19 Genomics UK (COG-UK) consortium An integrated national scale SARS-CoV-2 genomic surveillance network. *Lancet Microbe* **1**, e99–e100 (2020).
37. Hodcroft, E. B. et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. Preprint at *medRxiv* <https://doi.org/10.1101/2020.10.25.20219063> (2020).
38. FCO. *Travel Advice against all non-essential travel: Foreign Secretary’s statement, 17 March 2020*. <https://www.gov.uk/government/news/travel-advice-foreign-secretary-statement-17-march-2020> (2020).
39. Statistics, O.F.N. *2011 ONS Census*. <https://www.cambridge.gov.uk/media/1170/census-2011-cambridge-data.pdf> (2011).
40. DHSC, S. *The R Value and Growth Rate in the UK*. <https://www.gov.uk/guidance/the-r-number-in-the-uk> (2020).
41. Crawford, D. C. & Williams, S. M. Global variation in sequencing impedes SARS-CoV-2 surveillance. *PLoS Genet.* **17**, e1009620 (2021).
42. PHE. *Staying Alert and Staying Safe From 4 July*. <https://publichealthmatters.blog.gov.uk/2020/07/03/staying-alert-and-staying-safe-from-4-july/> (2020).
43. P. M. O. & T. R. H. B. J. *Prime Minister’s statement to the House on COVID-19: 23 June 2020*. <https://www.gov.uk/government/speeches/prime-ministers-statement-to-the-house-on-covid-19-23-june-2020> (2020).
44. CDC. *Science Brief: Options to Reduce Quarantine for Contacts of Persons with SARS-CoV-2 Infection Using Symptom Monitoring and Diagnostic Testing*. <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-options-to-reduce-quarantine.html> (2020).
45. Quick, J. *nCoV-2019 sequencing protocol v3 (LoCost) V.3*. <https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye> (2020).
46. Bonsall, D. et al. A comprehensive genomics solution for HIV surveillance and clinical monitoring in low-income settings. *J. Clin. Microbiol.* <https://doi.org/10.1128/JCM.00382-20> (2020).
47. Nicholls, S. M. et al. CLIMB-COVID: continuous integration supporting decentralised sequencing for SARS-CoV-2 genomic surveillance. *Genome Biol* **22**, 196 (2021).
48. Connor, T. R. et al. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. *Micro. Genom.* **2**, e000086 (2016).
49. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
50. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
51. Brooks, M. E. et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R. J.* **9**, 378 (2017).
52. Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. Emmeans: estimated marginal means, aka least-squares means. *R. package version 1*, 3 (2018).
53. Team, R. C. R: A language and environment for statistical computing. (2013).
54. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data —from vision to reality. *Euro Surveill.* <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494> (2017).
55. Cochrane, G., Karsch-Mizrachi, I. & Takagi, T., International Nucleotide Sequence Database, C. The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Res.* **44**, D48–D50 (2016).
56. Aggarwal, D. et al. *Genomic assessment of quarantine measures to prevent SARS-CoV-2 importation and transmission*. Github <https://github.com/COG-UK/travel-quarantine/> (2021).
57. Griffiths, E. J. T. et al. The PHA4GE SARS-CoV-2 Contextual Data Specification for Open Genomic Epidemiology. Preprints at <https://doi.org/10.20944/preprints202008.0220.v1> (2020).
58. UK, G. *Coronavirus (COVID-19): Travel Corridors*. <https://www.gov.uk/guidance/coronavirus-covid-19-travel-corridors>.

Acknowledgements

We thank members of the COVID-19 Genomics Consortium UK and Test and Trace contact tracers for their contributions to generating the data used in this study. We thank the Sanger Covid Team for assisting with Samples and Logistics. We thank Sarah Mitchell from the Department of Plant Sciences, University of Cambridge, for direction on statistical analysis. D.A. is a Clinical PhD Fellow and gratefully supported by the Wellcome Trust (Grant no. 222903/Z/21/Z). E.M.H. is supported by a UK Research and Innovation (UKRI) Fellowship: MR/S00291X/1. A.J.P., T.L.V. and G.M.S. gratefully acknowledge the support of

the Biotechnology and Biological Sciences Research Council (BBSRC); their research was funded by the BBSRC Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent project BBS/E/F/000PR10352, also Quadram Institute Bioscience BBSRC funded Core Capability Grant (project number BB/CCG1860/1). The COVID-19 Genomics UK (COG-UK) Consortium is supported by funding from the Medical Research Council (MRC) part of UK Research & Innovation (UKRI), the National Institute of Health Research (NIHR) and Genome Research Limited, operating as the Wellcome Sanger Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. For the purpose of open access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

Conceptualisation: M.C., E.M.H., S.J.P. and D.A.; Methodology: D.A., U.S., R.M., E.V., M.C. and E.M.H.; Supervision: M.C., E.M.H. and S.J.P.; Writing—original draft preparation: D.A., A.J.P. and G.M.S.; Writing—review and editing: All authors; Investigation: D.A., U.S., A.J.P., R.M., N.E., T.L.V. and N.M.T.; Formal analysis: D.A., U.S., R.M. and G.M.S.; Data curation: D.A., S.P., N.G., E.G., G.J.H., C.C., C.T., S.L. and A.H.; Funding acquisition: S.J.P. and E.M.H.; Resources and validation: G.J.H., C.C., C.T., S.L., A.H., D.A., A.J.P. and U.S.; Resources: COG-UK consortium.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28371-z>.

Correspondence and requests for materials should be addressed to Dinesh Aggarwal or Ewan M. Harrison.

Peer review information *Nature Communications* thanks Colin Simpson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

The COVID-19 Genomics UK (COG-UK) Consortium

Cherian Koshy¹¹, Amy Ash¹¹, Emma Wise¹², Nathan Moore¹², Matilde Mori¹², Nick Cortes¹², Jessica Lynch¹², Stephen Kidd¹², Derek J. Fairley¹³, Tanya Curran¹³, James P. McKenna¹³, Helen Adams¹⁴, Christophe Fraser¹⁵, Tanya Golubchik¹⁵, David Bonsall¹⁵, Mohammed O. Hassan-Ibrahim¹⁶, Cassandra S. Malone¹⁶, Benjamin J. Cogger¹⁶, Michelle Wantoch¹⁷, Nicola Reynolds¹⁷, Ben Warne³, Joshua Maksimovic¹⁸, Karla Spellman¹⁸, Kathryn McCluggage¹⁸, Michaela John¹⁸, Robert Beer¹⁸, Safiah Afifi¹⁸, Sian Morgan¹⁸, Angela Marchbank¹⁹, Anna Price¹⁹, Christine Kitchen¹⁹, Huw Gulliver¹⁹, Ian Merrick¹⁹, Joel Southgate¹⁹, Martyn Guest¹⁹, Robert Munn¹⁹, Trudy Workman¹⁹, Thomas R. Connor¹⁹, William Fuller¹⁹, Catherine Bresner¹⁹, Luke B. Snell²⁰, Amita Patel²⁰, Themoula Charalampous²¹, Gaia Nebbia²¹, Rahul Batra²¹, Jonathan Edgeworth²¹, Samuel C. Robson²², Angela H. Beckett²², David M. Aanensen²³, Anthony P. Underwood²³, Corin A. Yeats²³, Khalil Abudahab²³, Ben E. W. Taylor²³, Mirko Menegazzo²³, Gemma Clark²⁴, Wendy Smith²⁴, Manjinder Khakh²⁴, Vicki M. Fleming²⁴, Michelle M. Lister²⁴, Hannah C. Howson-Wells²⁴, Louise Berry²⁴, Tim Boswell²⁴, Amelia Joseph²⁴, Iona Willingham²⁴, Carl Jones²⁴, Christopher Holmes²⁵, Paul Bird²⁵, Thomas Helmer²⁵, Karlie Fallon²⁵, Julian Tang²⁵, Veena Raviprakash²⁶, Sharon Campbell²⁶, Nicola Sheriff²⁶, Victoria Blakey²⁶, Lesley-Anne Williams²⁶, Matthew W. Loose²⁷, Nadine Holmes²⁷, Christopher Moore²⁷, Matthew Carlile²⁷, Victoria Wright²⁷, Fei Sang²⁷, Johnny Debebe²⁷, Francesc Coll²⁸, Adrian W. Signell²⁹, Gilberto Betancor²⁹, Harry D. Wilson²⁹, Sahar Eldirdiri³⁰, Anita Kenyon³⁰, Thomas Davis³⁰, Oliver G. Pybus³¹, Louis du Plessis³¹, Alex E. Zarebski³¹, Jayna Raghwani³¹, Moritz U. G. Kraemer³¹, Sarah Francois³¹, Stephen W. Attwood³¹, Tetyana I. Vasylyeva³¹, Marina Escalera Zamudio³¹, Bernardo Gutierrez³¹, M. Estee Torok³², William L. Hamilton³², Ian G. Goodfellow³³, Grant Hall³³, Aminu S. Jahun³³, Yasmin Chaudhry³³, Myra Hosmillo³³, Malte L. Pinckert³³, Iliana Georgana³³, Samuel Moses³⁴, Hannah Lowe³⁴, Luke Bedford³⁵, Jonathan Moore³⁶, Susanne Stonehouse³⁶, Chloe L. Fisher³⁷, Ali R. Awan³⁷, John BoYes³⁸, Judith Breuer³⁹, Kathryn A. Harris³⁹, Julianne R. Brown³⁹, Divya Shah³⁹, Laura Atkinson³⁹, Jack C. D. Lee³⁹,

Nathaniel Storey³⁹, Flavia Flaviani⁴⁰, Adela Alcolea-Medina⁴¹, Rebecca Williams⁴², Gabrielle Vernet⁴², Michael R. Chapman⁴³, Lisa J. Levett⁴⁴, Judith Heaney⁴⁴, Wendy Chatterton⁴⁴, Monika Pusok⁴⁴, Li Xu-McCrae⁴⁵, Darren L. Smith⁴⁶, Matthew Bashton⁴⁶, Gregory R. Young⁴⁶, Alison Holmes⁴⁷, Paul A. Randell⁴⁷, Alison Cox⁴⁷, Pinglawathee Madona⁴⁷, Frances Bolt⁴⁷, James Price⁴⁷, Siddharth Mookerjee⁴⁷, Manon Ragonnet-Cronin⁶, Fabricia F. Nascimento⁶, David Jorgensen⁶, Igor Siveroni⁶, Rob Johnson⁶, Olivia Boyd⁶, Lily Geidelberg⁶, Erik M. Volz⁶, Aileen Rowan⁶, Graham P. Taylor⁶, Katherine L. Smollett⁴⁸, Nicholas J. Loman⁴⁹, Joshua Quick⁴⁹, Claire McMurray⁴⁹, Joanne Stockton⁴⁹, Sam Nicholls⁴⁹, Will Rowe⁴⁹, Radoslaw Poplawski⁴⁹, Alan McNally⁴⁹, Rocio T. Martinez Nunez⁵⁰, Jenifer Mason⁵¹, Trevor I. Robinson⁵¹, Elaine O'Toole⁵¹, Joanne Watts⁵¹, Cassie Breen⁵¹, Angela Cowell⁵¹, Graciela Sluga⁵², Nicholas W. Machin⁵³, Shazaad S. Y. Ahmad⁵³, Ryan P. George⁵³, Fenella Halstead⁵⁴, Venkat Sivaprakasam⁵⁴, Wendy Hogsden⁵⁴, Chris J. Illingworth⁵⁵, Chris Jackson⁵⁵, Emma C. Thomson⁵⁶, James G. Shepherd⁵⁶, Patawee Asamaphan⁵⁶, Marc O. Niebel⁵⁶, Kathy K. Li⁵⁶, Rajiv N. Shah⁵⁶, Natasha G. Jesudason⁵⁶, Lily Tong⁵⁶, Alice Broos⁵⁶, Daniel Mair⁵⁶, Jenna Nichols⁵⁶, Stephen N. Carmichael⁵⁶, Kyriaki Nomikou⁵⁶, Elihu Aranday-Cortes⁵⁶, Natasha Johnson⁵⁶, Igor Starinskij⁵⁶, Ana da Silva Filipe⁵⁶, David L. Robertson⁵⁶, Richard J. Orton⁵⁶, Joseph Hughes⁵⁶, Sreenu Vattipally⁵⁶, Joshua B. Singer⁵⁶, Seema Nickbakhsh⁵⁶, Antony D. Hale⁵⁷, Louissa R. Macfarlane-Smith⁵⁷, Katherine L. Harper⁵⁷, Holli Carden⁵⁷, Yusri Taha⁵⁸, Brendan A. I. Payne⁵⁸, Shirelle Burton-Fanning⁵⁸, Sheila Waugh⁵⁸, Jennifer Collins⁵⁸, Gary Eltringham⁵⁸, Steven Rushton⁵⁹, Sarah O'Brien⁵⁹, Amanda Bradley⁶⁰, Alasdair Maclean⁶⁰, Guy Mollett⁶⁰, Rachel Blacow⁶⁰, Kate E. Templeton⁶¹, Martin P. McHugh⁶¹, Rebecca Dewar⁶¹, Elizabeth Wastenge⁶¹, Samir Dervisevic⁶², Rachael Stanley⁶², Emma J. Meader⁶², Lindsay Coupland⁶², Louise Smith⁶³, Clive Graham⁶⁴, Edward Barton⁶⁴, Debra Padgett⁶⁴, Garren Scott⁶⁴, Emma Swindells⁶⁵, Jane Greenaway⁶⁵, Andrew Nelson⁶⁶, Clare M. McCann⁶⁶, Wen C. Yew⁶⁶, Monique Andersson⁶⁷, Timothy Peto⁶⁷, Anita Justice⁶⁷, David Eyre⁶⁷, Derrick Crook⁶⁷, Tim J. Sloan⁶⁸, Nichola Duckworth⁶⁸, Sarah Walsh⁶⁸, Anoop J. Chauhan⁶⁹, Sharon Glaysher⁶⁹, Kelly Bicknell⁶⁹, Sarah Wyllie⁶⁹, Scott Elliott⁶⁹, Allyson Lloyd⁶⁹, Robert Impey⁶⁹, Nick Levene⁷⁰, Lynn Monaghan⁷⁰, Declan T. Bradley⁷¹, Tim Wyatt⁷¹, Elias Allara⁷², Clare Pearson⁷², Husam Osman⁷², Andrew Bosworth⁷², Esther Robinson⁷², Peter Muir⁷², Ian B. Vipond⁷², Richard Hopes⁷², Hannah M. Pymont⁷², Stephanie Hutchings⁷², Martin D. Curran⁷³, Surendra Parmar⁷³, Angie Lackenby², Tamyo Mbisa², Steven Platt², Shahjahan Miah², David Bibby², Carmen Manso², Jonathan Hubb², Meera Chand^{2,105}, Gavin Dabrera², Mary Ramsay², Daniel Bradshaw², Alicia Thornton², Richard Myers², Ulf Schaefer^{2,104}, Natalie Groves², Eileen Gallagher², David Lee², David Williams², Nicholas Ellaby², Ian Harrison², Hassan Hartman², Nikos Manesis², Vineet Patel², Chloe Bishop², Vicki Chalker², Juan Ledesma², Katherine A. Twohig², Matthew T. G. Holden⁷⁴, Sharif Shaaban⁷⁴, Alec Birchley⁷⁵, Alexander Adams⁷⁵, Alisha Davies⁷⁵, Amy Gaskin⁷⁵, Amy Plimmer⁷⁵, Bree Gatica-Wilcox⁷⁵, Caoimhe McKerr⁷⁵, Catherine Moore⁷⁵, Chris Williams⁷⁵, David Heyburn⁷⁵, Elen De Lacy⁷⁵, Ember Hilvers⁷⁵, Fatima Downing⁷⁵, Giri Shankar⁷⁵, Hannah Jones⁷⁵, Hibo Asad⁷⁵, Jason Coombes⁷⁵, Joanne Watkins⁷⁵, Johnathan M. Evans⁷⁵, Laia Fina⁷⁵, Laura Gifford⁷⁵, Lauren Gilbert⁷⁵, Lee Graham⁷⁵, Malorie Perry⁷⁵, Mari Morgan⁷⁵, Matthew Bull⁷⁵, Michelle Cronin⁷⁵, Nicole Pacchiarini⁷⁵, Noel Craine⁷⁵, Rachel Jones⁷⁵, Robin Howe⁷⁵, Sally Corden⁷⁵, Sara Rey⁷⁵, Sara Kumziene-SummerhaYes⁷⁵, Sarah Taylor⁷⁵, Simon Cottrell⁷⁵, Sophie Jones⁷⁵, Sue Edwards⁷⁵, Justin O'Grady⁵, Andrew J. Page ^{5,104}, Alison E. Mather⁵, David J. Baker⁵, Steven Rudder⁵, Alp Aydin⁵, Gemma L. Kay⁵, Alexander J. Trotter⁵, Nabil-Fareed Alikhan⁵, Leonardo de Oliveira Martins⁵, Thanh Le-Viet⁵, Lizzie Meadows⁵, Anna Casey⁷⁶, Liz Ratcliffe⁷⁶, David A. Simpson⁷⁷, Zoltan Molnar⁷⁷, Thomas Thompson⁷⁷, Erwan Acheson⁷⁷, Jane A. H. Masoli⁷⁸, Bridget A. Knight⁷⁸, Sian Ellard⁷⁸, Cressida Auckland⁷⁸, Christopher R. Jones⁷⁸, Tabitha W. Mahungu⁷⁹, Dianne Irish-Tavares⁷⁹, Tanzina Haque⁷⁹, Jennifer Hart⁷⁹, Eric Witele⁷⁹, Melisa L. Fenton⁸⁰, Ashok Dadrah⁸⁰,

Amanda Symmonds⁸⁰, Tranpriti Saluja⁸⁰, Yann Bourgeois⁸¹, Garry P. Scarlett⁸¹, Katie F. Loveson⁸², Salman Goudarzi⁸², Christopher Fearn⁸², Kate Cook⁸², Hannah Dent⁸², Hannah Paul⁸², David G. Partridge⁸³, Mohammad Raza⁸³, Cariad Evans⁸³, Kate Johnson⁸³, Steven Liggett⁸⁴, Paul Baker⁸⁴, Stephen Bonner⁸⁴, Sarah Essex⁸⁴, Ronan A. Lyons⁸⁵, Kordo Saeed⁸⁶, Adhyana I. K. Mahanama⁸⁶, Buddhini Samaraweera⁸⁶, Siona Silveira⁸⁶, Emanuela Pelosi⁸⁶, Eleri Wilson-Davies⁸⁶, Rachel J. Williams⁸⁷, Mark Kristiansen⁸⁷, Sunando Roy⁸⁷, Charlotte A. Williams⁸⁷, Marius Cotic⁸⁷, Nadua Bayzid⁸⁷, Adam P. Westhorpe⁸⁷, John A. Hartley⁸⁷, Riaz Jannoo⁸⁷, Helen L. Lowe⁸⁷, Angeliki Karamani⁸⁷, Leah Ensell⁸⁷, Jacqui A. Prieto⁸⁸, Sarah Jeremiah⁸⁸, Dimitris Grammatopoulos⁸⁹, Sarojini Pandey⁸⁹, Lisa Berry⁸⁹, Katie Jones⁸⁹, Alex Richter⁹⁰, Andrew Beggs⁹⁰, Angus Best⁹¹, Benita Percival⁹¹, Jeremy Mirza⁹¹, Oliver Megram⁹¹, Megan Mayhew⁹¹, Liam Crawford⁹¹, Fiona Ashcroft⁹¹, Emma Moles-Garcia⁹¹, Nicola Cumley⁹¹, Colin P. Smith⁹², Giselda Bucca⁹², Andrew R. Hesketh⁹², Beth Blane⁹³, Sophia T. Girgis⁹³, Danielle Leek⁹³, Sushmita Sridhar⁹³, Sally Forrest⁹³, Claire Cormie⁹³, Harmeet K. Gill⁹³, Joana Dias⁹³, Ellen E. Higginson⁹³, Mailis Maes⁹³, Jamie Young⁹³, Leanne M. Kermack⁹³, Ravi K. Gupta⁹³, Catherine Ludden⁹³, Sharon J. Peacock⁹³, Sophie Palmer⁹³, Carol M. Churcher⁹³, Nazreen F. Hadjirin⁹³, Alessandro M. Carabelli⁹³, Ellena Brooks⁹³, Kim S. Smith⁹³, Katerina Galai⁹³, Georgina M. McManus⁹³, Chris Ruis⁹³, Rose K. Davidson⁹⁴, Andrew Rambaut⁹⁵, Thomas Williams⁹⁵, Carlos E. Balcazar⁹⁵, Michael D. Gallagher⁹⁵, Áine O'Toole⁹⁵, Stefan Rooke⁹⁵, Verity Hill⁹⁵, Kathleen A. Williamson⁹⁵, Thomas D. Stanton⁹⁵, Stephen L. Michell⁹⁶, Claire M. Bewshea⁹⁶, Ben Temperton⁹⁶, Michelle L. Michelsen⁹⁶, Joanna Warwick-Dugdale⁹⁶, Robin Manley⁹⁶, Audrey Farbos⁹⁶, James W. Harrison⁹⁶, Christine M. Sambles⁹⁶, David J. Studholme⁹⁶, Aaron R. Jeffries⁹⁶, Alistair C. Darby⁹⁷, Julian A. Hiscox⁹⁷, Steve Paterson⁹⁷, Miren Iturriza-Gomara⁹⁷, Kathryn A. Jackson⁹⁷, Anita O. Lucaci⁹⁷, Edith E. Vamos⁹⁷, Margaret Hughes⁹⁷, Lucille Rainbow⁹⁷, Richard Eccles⁹⁷, Charlotte Nelson⁹⁷, Mark Whitehead⁹⁷, Lance Turtle⁹⁷, Sam T. Haldenby⁹⁷, Richard Gregory⁹⁷, Matthew Gemmell⁹⁷, Claudia Wierzbicki⁹⁷, Hermione J. Webster⁹⁷, Thushan I. de Silva⁹⁸, Nikki Smith⁹⁸, Adrienn Angyal⁹⁸, Benjamin B. Lindsey⁹⁸, Danielle C. Groves⁹⁸, Luke R. Green⁹⁸, Dennis Wang⁹⁸, Timothy M. Freeman⁹⁸, Matthew D. Parker⁹⁸, Alexander J. Keeley⁹⁸, Paul J. Parsons⁹⁸, Rachel M. Tucker⁹⁸, Rebecca Brown⁹⁸, Matthew Wyles⁹⁸, Max Whiteley⁹⁸, Peijun Zhang⁹⁸, Marta Gallis⁹⁸, Stavroula F. Louka⁹⁸, Chrystala Constantinidou⁹⁹, Meera Unnikrishnan⁹⁹, Sascha Ott⁹⁹, Jeffrey K. J. Cheng⁹⁹, Hannah E. Bridgewater⁹⁹, Lucy R. Frost⁹⁹, Grace Taylor-Joyce⁹⁹, Richard Stark⁹⁹, Laura Baxter⁹⁹, Mohammad T. Alam⁹⁹, Paul E. Brown⁹⁹, Dinesh Aggarwal⁹⁹, Alberto C. Cerda¹⁰⁰, Tammy V. Merrill¹⁰⁰, Rebekah E. Wilson¹⁰⁰, Patrick C. McClure¹⁰¹, Joseph G. Chappell¹⁰¹, Theocharis Tsoleridis¹⁰¹, Jonathan Ball¹⁰¹, David Buck¹⁰², John A. Todd¹⁰², Angie Green¹⁰², Amy Trebes¹⁰², George MacIntyre-Cockett¹⁰², Mariateresa de Cesare¹⁰², Alex Alderton⁴, Roberto Amato⁴, Cristina V. Ariani⁴, Mathew A. Beale⁴, Charlotte Beaver⁴, Katherine L. Bellis⁴, Emma Betteridge⁴, James Bonfield⁴, John Danesh⁴, Matthew J. Dorman⁴, Eleanor Drury⁴, Ben W. Farr⁴, Luke Foulser⁴, Sonia Goncalves⁴, Scott Goodwin⁴, Marina Gourtovaia⁴, Ewan M. Harrison⁴, David K. Jackson⁴, Dorota Jamroz⁴, Ian Johnston⁴, Leanne Kane⁴, Sally Kay⁴, Jon-Paul Keatley⁴, Dominic Kwiatkowski⁴, Cordelia F. Langford⁴, Mara Lawniczak⁴, Laura Letchford⁴, Rich Livett⁴, Stephanie Lo⁴, Inigo Martincorena⁴, Samantha McGuigan⁴, Rachel Nelson⁴, Steve Palmer⁴, Naomi R. Park⁴, Minal Patel⁴, Liam Prestwood⁴, Christoph Puethe⁴, Michael A. Quail⁴, Shavanthi Rajatileka⁴, Carol Scott⁴, Lesley Shirley⁴, John Sillitoe⁴, Michael H. Spencer Chapman⁴, Scott A. J. Thurston⁴, Gerry Tonkin-Hill⁴, Danni Weldon⁴, Diana Rajan⁴, Iraad F. Bronner⁴, Louise Aigrain⁴, Nicholas M. Redshaw⁴, Stefanie V. Lensing⁴, Robert Davies⁴, Andrew Whitwham⁴, Jennifer Liddle⁴, Kevin Lewis⁴, Jaime M. Tovar-Corona⁴, Steven Leonard⁴, Jillian Durham⁴, Andrew R. Bassett⁴, Shane McCarthy⁴, Robin J. Moll⁴, Keith James⁴, Karen Oliver⁴, Alex Makunin⁴, Jeff Barrett⁴ & Rory N. Gunson¹⁰³

¹¹Barking, Havering and Redbridge University Hospitals NHS Trust, Romford, UK. ¹²Basingstoke Hospital, Basingstoke, UK. ¹³Belfast Health & Social Care Trust, Belfast, UK. ¹⁴Betsi Cadwaladr University Health Board, Bangor, UK. ¹⁵Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ¹⁶Brighton and Sussex University Hospitals NHS Trust, Brighton, UK. ¹⁷Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. ¹⁸Cardiff and Vale University Health Board, Cardiff, UK. ¹⁹Cardiff University, Cardiff, UK. ²⁰Centre for Clinical Infection & Diagnostics Research, St. Thomas' Hospital and Kings College London, London, UK. ²¹Centre for Clinical Infection and Diagnostics Research, Department of Infectious Diseases, Guy's and St Thomas' NHS Foundation Trust, London, UK. ²²Centre for Enzyme Innovation, University of Portsmouth (PORT), Portsmouth, UK. ²³Centre for Genomic Pathogen Surveillance, University of Oxford, Oxford, UK. ²⁴Clinical Microbiology Department, Queens Medical Centre, Nottingham, UK. ²⁵Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester, UK. ²⁶County Durham and Darlington NHS Foundation Trust, Darlington, UK. ²⁷Deep Seq, School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK. ²⁸Department of Infection Biology, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK. ²⁹Department of Infectious Diseases, King's College London, London, UK. ³⁰Department of Microbiology, Kettering General Hospital, Kettering, UK. ³¹Department of Zoology, University of Oxford, Oxford, UK. ³²Departments of Infectious Diseases and Microbiology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ³³Division of Virology, Department of Pathology, University of Cambridge, Cambridge, UK. ³⁴East Kent Hospitals University NHS Foundation Trust, Canterbury, UK. ³⁵East Suffolk and North Essex NHS Foundation Trust, Colchester, UK. ³⁶Gateshead Health NHS Foundation Trust, Gateshead, UK. ³⁷Genomics Innovation Unit, Guy's and St. Thomas' NHS Foundation Trust, London, UK. ³⁸Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK. ³⁹Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. ⁴⁰Guy's and St. Thomas' BRC, London, UK. ⁴¹Guy's and St. Thomas' Hospitals, London, UK. ⁴²Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK. ⁴³Health Data Research UK Cambridge, Cambridge, UK. ⁴⁴Health Services Laboratories, London, UK. ⁴⁵Heartlands Hospital, Birmingham, UK. ⁴⁶Hub for Biotechnology in the Built Environment, Northumbria University, Newcastle, UK. ⁴⁷Imperial College Hospitals NHS Trust, London, UK. ⁴⁸Institute of Biodiversity, Animal Health & Comparative Medicine, Glasgow, UK. ⁴⁹Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK. ⁵⁰King's College London, London, UK. ⁵¹Liverpool Clinical Laboratories, Liverpool, UK. ⁵²Maidstone and Tunbridge Wells NHS Trust, Tunbridge Wells, UK. ⁵³Manchester University NHS Foundation Trust, Manchester, UK. ⁵⁴Microbiology Department, Wye Valley NHS Trust, Hereford, UK. ⁵⁵MRC Biostatistics Unit, University of Cambridge, Cambridge, UK. ⁵⁶MRC-University of Glasgow Centre for Virus Research, Glasgow, UK. ⁵⁷National Infection Service, PHE and Leeds Teaching Hospitals Trust, Leeds, UK. ⁵⁸Newcastle Hospitals NHS Foundation Trust, Newcastle, UK. ⁵⁹Newcastle University, Newcastle, UK. ⁶⁰NHS Greater Glasgow and Clyde, Glasgow, UK. ⁶¹NHS Lothian, Edinburgh, UK. ⁶²Norfolk and Norwich University Hospital, Norwich, UK. ⁶³Norfolk County Council, Norwich, UK. ⁶⁴North Cumbria Integrated Care NHS Foundation Trust, Carlisle, UK. ⁶⁵North Tees and Hartlepool NHS Foundation Trust, Stockton on Tees, UK. ⁶⁶Northumbria University, Newcastle, UK. ⁶⁷Oxford University Hospitals NHS Foundation Trust, Oxford, UK. ⁶⁸PathLinks, Northern Lincolnshire & Goole NHS Foundation Trust, Grimsby, UK. ⁶⁹Portsmouth Hospitals University NHS Trust, Portsmouth, UK. ⁷⁰Princess Alexandra Hospital Microbiology Dept, Harlow, UK. ⁷¹Public Health Agency, Belfast, UK. ⁷²Public Health England, London, UK. ⁷³Public Health England, Clinical Microbiology and Public Health Laboratory, Cambridge, UK. ⁷⁴Public Health Scotland, Edinburgh, UK. ⁷⁵Public Health Wales NHS Trust, Cardiff, UK. ⁷⁶Queen Elizabeth Hospital, London, UK. ⁷⁷Queen's University Belfast, Belfast, UK. ⁷⁸Royal Devon and Exeter NHS Foundation Trust, Exeter, UK. ⁷⁹Royal Free NHS Trust, London, UK. ⁸⁰Sandwell and West Birmingham NHS Trust, Birmingham, UK. ⁸¹School of Biological Sciences, University of Portsmouth (PORT), Portsmouth, UK. ⁸²School of Pharmacy and Biomedical Sciences, University of Portsmouth (PORT), Portsmouth, UK. ⁸³Sheffield Teaching Hospitals, Sheffield, UK. ⁸⁴South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK. ⁸⁵Swansea University, Swansea, UK. ⁸⁶University Hospitals Southampton NHS Foundation Trust, Southampton, UK. ⁸⁷University College London, London, UK. ⁸⁸University Hospital Southampton NHS Foundation Trust, Southampton, UK. ⁸⁹University Hospitals Coventry and Warwickshire, Coventry, UK. ⁹⁰University of Birmingham, Birmingham, UK. ⁹¹University of Birmingham Turnkey Laboratory, Birmingham, UK. ⁹²University of Brighton, Brighton, UK. ⁹³University of Cambridge, Cambridge, UK. ⁹⁴University of East Anglia, Norwich, UK. ⁹⁵University of Edinburgh, Edinburgh, UK. ⁹⁶University of Exeter, Exeter, UK. ⁹⁷University of Liverpool, Liverpool, UK. ⁹⁸University of Sheffield, Sheffield, UK. ⁹⁹University of Warwick, Warwick, UK. ¹⁰⁰Viapath, Guy's and St Thomas' NHS Foundation Trust, and King's College Hospital NHS Foundation Trust, London, UK. ¹⁰¹Virology, School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK. ¹⁰²Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ¹⁰³West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, Glasgow, UK.