

Northumbria Research Link

Citation: Nickbakhsh, Sema, Hughes, Joseph, Christofidis, Nicolaos, Griffiths, Emily, Shaaban, Sharif, Enright, Jessica, Smollett, Katherine, Nomikou, Kyriaki, Palmalux, Natasha, Tong, Lily, Carmichael, Stephen, Sreenu, Vattipally B., Orton, Richard, Goldstein, Emily J., Tomb, Rachael M., The COVID-19 Genomics UK (COG-UK) Consortium, , Templeton, Kate, Gunson, Rory N., da Silva Filipe, Ana, Milosevic, Catriona, Thomson, Emma, Robertson, David L., Holden, Matthew T.G., Illingworth, Christopher J.R., Smith-Palmer, Alison, Smith, Darren, Bashton, Matthew, Young, Greg, McCann, Clare, Nelson, Andrew, Crown, Matthew, Henderson, John and Yew, Wen Chyin (2022) Genomic epidemiology of SARS-CoV-2 in a university outbreak setting and implications for public health planning. *Scientific Reports*, 12 (1). p. 11735. ISSN 2045-2322

Published by: Nature Publishing Group

URL: <https://doi.org/10.1038/s41598-022-15661-1> <<https://doi.org/10.1038/s41598-022-15661-1>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/49744/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary



OPEN

Genomic epidemiology of SARS-CoV-2 in a university outbreak setting and implications for public health planning

Sema Nickbakhsh^{1,2,125}✉, Joseph Hughes^{1,3,125}, Nicolaos Christofidis¹, Emily Griffiths¹, Sharif Shaaban¹, Jessica Enright⁴, Katherine Smollett³, Kyriaki Nomikou³, Natasha Palmalux³, Lily Tong³, Stephen Carmichael³, Vattipally B. Sreenu³, Richard Orton³, Emily J. Goldstein⁵, Rachael M. Tomb⁵, The COVID-19 Genomics UK (COG-UK) Consortium^{6*}, Kate Templeton⁷, Rory N. Gunson⁵, Ana da Silva Filipe³, Catriona Milosevic⁸, Emma Thomson^{1,3}, David L. Robertson^{1,3}, Matthew T. G. Holden^{1,9}, Christopher J. R. Illingworth^{3,10,11,126}✉ & Alison Smith-Palmer^{1,126}

Whole genome sequencing of SARS-CoV-2 has occurred at an unprecedented scale, and can be exploited for characterising outbreak risks at the fine-scale needed to inform control strategies. One setting at continued risk of COVID-19 outbreaks are higher education institutions, associated with student movements at the start of term, close living conditions within residential halls, and high social contact rates. Here we analysed SARS-CoV-2 whole genome sequences in combination with epidemiological data to investigate a large cluster of student cases associated with University of Glasgow accommodation in autumn 2020, Scotland. We identified 519 student cases of SARS-CoV-2 infection associated with this large cluster through contact tracing data, with 30% sequencing coverage for further analysis. We estimated at least 11 independent introductions of SARS-CoV-2 into the student population, with four comprising the majority of detected cases and consistent with separate outbreaks. These four outbreaks were curtailed within a week following implementation of control measures. The impact of student infections on the local community was short-term despite an underlying increase in community infections. Our study highlights the need for context-specific information in the formation of public health policy for higher educational settings.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019 (COVID-19), is persisting endemically in the UK with long-term implications for public health planning and preparedness. As at 2nd June 2022, 73% of the Scottish population aged 12 and over have received three doses of COVID-19 vaccine¹. However, lower population immunity among children, adolescents and younger adults², the occurrence of infection despite vaccination³, and the emergence of vaccine escape variants^{4,5} threaten the long-term ability of the national vaccination programme to suppress the spread of infection and prevent disease.

¹Public Health Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, UK. ²Institute of Biodiversity, Animal Health & Comparative Medicine, University of Glasgow, Graham Kerr Building, Glasgow G12 8QQ, UK. ³MRC-University of Glasgow Centre for Virus Research, 464 Bearsden Road, Glasgow G61 1QH, UK. ⁴School of Computing Science, University of Glasgow, 18 Lilybank Gardens, Glasgow G12 8RZ, UK. ⁵West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, Glasgow Royal Infirmary, New Lister Building, Glasgow G31 2ER, UK. ⁶Royal Infirmary of Edinburgh, NHS Lothian, 51 Little France Crescent, Edinburgh EH16 4SA, UK. ⁷NHS Greater Glasgow and Clyde, Gartnavel General Hospital, 1055 Great Western Road, Glasgow G12 0XH, UK. ⁸School of Medicine, University of St Andrews, North Haugh, St Andrews KY16 9TF, UK. ⁹Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK. ¹⁰MRC Biostatistics Unit, University of Cambridge, East Forvie Building, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK. ¹²⁵These authors contributed equally: Sema Nickbakhsh and Joseph Hughes. ¹²⁶These authors jointly supervised this work: Christopher J. R. Illingworth and Alison Smith-Palmer. *A list of authors and their affiliations appears at the end of the paper. ✉email: Sema.Nickbakhsh@phs.scot; Christopher.Illingworth@glasgow.ac.uk

With the exception of the first SARS-CoV-2 epidemic wave in 2020, cases among children, adolescents, and young adults have represented a high fraction of confirmed cases in Scotland. Younger age groups may contribute to seasonal patterns of COVID-19 in the long-term, with annual outbreaks in educational settings a well-recognised feature of acute respiratory infections in temperate climates⁶. Although we note that large and frequent COVID-19 clusters are also associated with settings involving older adults such as leisure, sporting events, and occupation^{7–9}.

With respect to higher education settings, the risk of respiratory infection outbreaks have been brought to the forefront by COVID-19, with the implementation of social distancing and remote learning to limit cases¹⁰. Universities present a unique epidemiological scenario for study, with the movement of students to and from domicile locations at the start and end of academic terms presenting a risk for introducing viruses into student populations and for spillover into local communities¹¹. Theoretical models suggest that SARS-CoV-2 will more easily transmit in denser populations depending on vaccination rates, adherence to social distancing, and effectiveness of contact tracing and isolation measures^{12,13}. With a potential increased likelihood of asymptomatic infection¹⁴ and lower vaccination coverage¹, student populations are at risk of seeding outbreaks among their social groups as well as exposing the local community and non-term-time contacts. Reports highlighted the potential for a large number of cases in higher education settings, particularly in association with densely populated residential halls^{15,16}.

Expansive SARS-CoV-2 genomic surveillance conducted in the UK, with over 300,000 whole genome sequences generated by June 2022 in Scotland alone, provides a means to characterise outbreak risks in a setting-specific manner. In combination with epidemiological data, genomic surveillance has enabled fine-grain characterisation of SARS-CoV-2 outbreaks in other semi-closed settings such as hospital wards^{17,18} and care homes¹⁹. To date one study in the UK has taken such an approach to understand SARS-CoV-2 outbreaks in a tertiary education setting¹⁵. This study identified persistent transmission among students, though with little evidence of transmission to the local community. However, the social and community contexts of universities in the UK differ, such that patterns of spread within a collegiate university, a single institution organised into self-governing colleges with associated accommodation, may not be directly applicable to universities with a more centralised organisational structure.

Here we use genomic epidemiology to investigate a large cluster of cases among University of Glasgow (UoG) students in Scotland, UK, arising in the autumn of 2020 at the start of the 2020/21 academic year²⁰. At this point reported case numbers were relatively low, although with the second epidemic wave in Scotland already underway²¹. Our study provides a genomically-informed insight into epidemiological risks associated with SARS-CoV-2 outbreaks in a university accommodation setting in which students were resident in halls integrated with the local community in the city of Glasgow.

Our study contributes to a comprehensive understanding of the role that residential accommodation can play in the spread of SARS-CoV-2, and highlights the success of rapidly implemented control measures on mitigating any subsequent impact on the wider university population and the local community. Given the current lower 3rd dose vaccine uptake among young adults in Scotland (as at June 2022)¹, and in the event of a novel SARS-CoV-2 variant of concern, there is a need for public health planning and policy formation to consider the higher educational setting, particularly in preparation for winter seasons. We propose that consideration is given to the varying accommodation, academic, and social contexts across student populations.

Results

Outbreak investigation. As of the 2020/21 academic year, the UoG hosts approximately 33,000 students from more than 140 countries, with 97% of undergraduate students aged 16–29 years, and with 58% of the 6100 new undergraduate students domiciled in Scotland²². The university provides multiple residential halls predominantly housing undergraduate students²³.

A cluster of COVID-19 student cases associated with two geographically distinct residential halls was identified by the UoG and the NHS Greater Glasgow and Clyde (NHSGGC) Public Health Protection Unit at the start of the 2020/21 academic year, immediately following Freshers' week (12th to 18th September) (Fig. 1a). The outbreak peaked and was subsequently curtailed within around seven days of being identified, following the implementation of non-pharmaceutical control measures commonly in use at the time for controlling SARS-CoV-2; including infection control, contact tracing, increased availability of testing, and support for students to enable self-isolation.

A total of 1039 student SARS-CoV-2 positive cases were identified, with 616 (59.3%) resident at one of six university residential halls spanning 13 postcode locations. The remaining 423 (40.7%) students were deemed to reside in private accommodation (see “Methods”). The median case age was 18 years (IQR 18–19 years; range 16–51 years), consistent with a majority undergraduate population, and 57.2% (n = 591) were female consistent with a known gender skew among the wider student population²⁴. Of the 616 student cases associated with residential halls, 519 (84.2%) had respiratory specimen dates from 19th September to 26th September 2020 spanning the residential halls cluster. Based on these data, the student cluster included around 9% of the new undergraduate population overall. SARS-CoV-2 whole genome sequences were available for 220 cases, with specimen dates spanning 5th September to 18th November 2020; 73.2% were from students residing in university residential halls with specimen dates from 19th September to 26th September 2020 (Fig. 1b). The median age among sequenced cases was 18 years (range 17–25 years, IQR 18–19 years) and 138 (62.7%) were female.

SARS-CoV-2 genetic diversity. Among the 220 student cases, multiple global lineages (Pango nomenclature)^{25,26} were detected (Fig. 1c), the most common being B.1 (n = 114, 52.8%), B.1.1.315 (n = 41, 19.0%), B.1.177

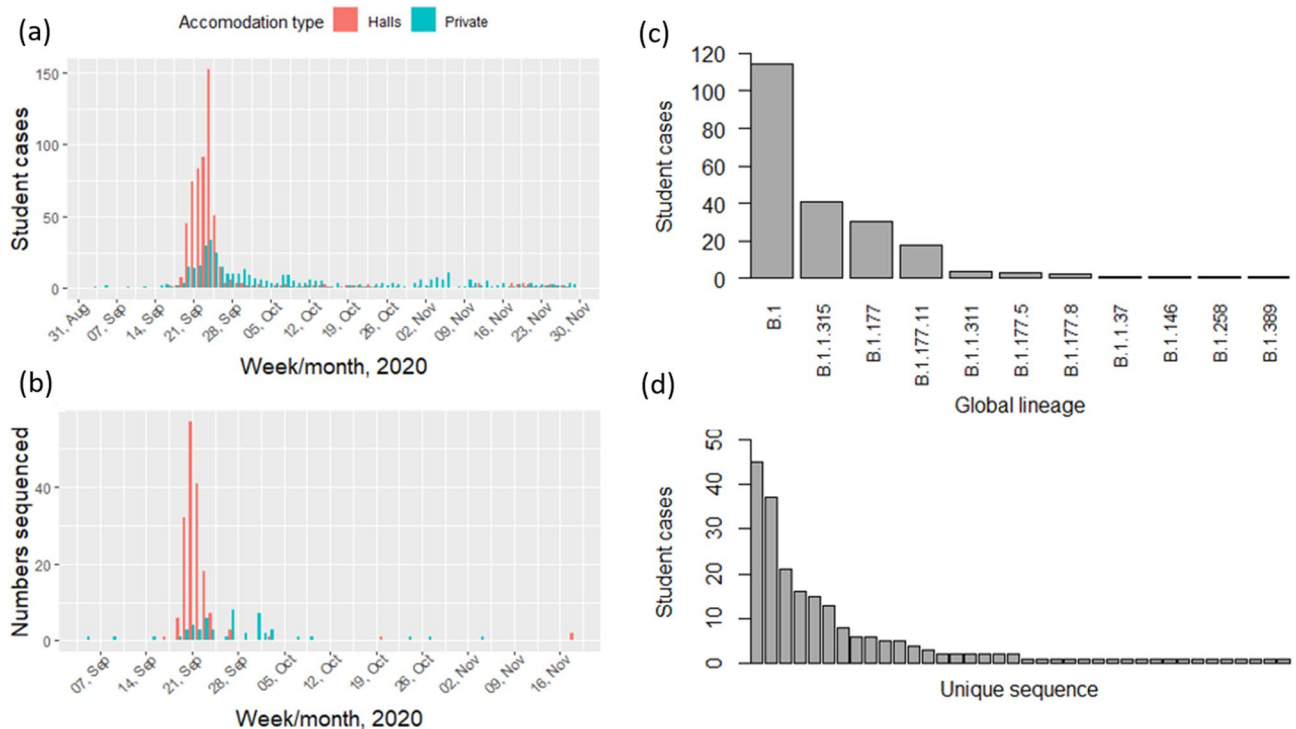


Figure 1. (a) Numbers of University of Glasgow student cases identified from contact tracing information over time, and (b) for the subset with sequenced samples, according to university residential halls or private accommodation. (c) Frequency distributions of SARS-CoV-2 lineages and (d) unique sequences identified among the student cases.

($n = 30$, 13.9%) and B.1.177.11 ($n = 18$, 8.3%). Seven further Pango lineages were detected at low numbers ($n = 13$, 5.9%). A total of thirty-nine distinct SARS-CoV-2 sequences were identified (Fig. 1d).

Putative transmission clusters were identified from 214 high quality SARS-CoV-2 genome sequences (see “Methods”) using the A2B-COVID software package¹⁷. We use the term ‘transmission cluster’ to denote a set of potentially linked cases consistent with having arisen from a single introduction into the student population (see “Methods”; Supplementary Fig. S1, supplementary material). Of the 214 high quality sequences, 157 were obtained from students resident in halls with specimen dates spanning 19th to 26th September 2020, providing 30% coverage of the student halls COVID-19 cluster. Four large transmission clusters were identified, comprising 191 sequences; mapping these onto a phylogenetic tree identified them with the four clades (highlighted in Fig. 2). Among the remaining 23 sequences, three further putative transmission clusters were found, each comprising two or three cases, alongside 16 sequences that were not positioned closely with other transmission clusters.

While each putative transmission cluster is consistent with describing a distinct introduction into the student population, a more conservative estimate was made that at least 11 distinct introductions of SARS-CoV-2 had occurred. To derive this, each transmission cluster and any sequences descended from it was counted as a single introduction. Further, the sequences intermediate to transmission clusters 2 and 3 were regarded as potentially arising from a single introduction (Fig. 2). In either case, the majority of introductions of SARS-CoV-2 into the university failed to cause substantial onward transmission among students.

Genomic evidence is consistent with introductions from within and outside Scotland. Determining the most recent common ancestor of the sequences associated with the four main transmission clusters positioned them within more expansive UK-wide phylogenetic clades circulating at that time. These clades combined both student and community sequences and revealed potentially distinct origins for the introductions to the University forming four main transmission clusters. Investigating the proportion of genome sequences falling into each of the four clades by regional location on a weekly basis, we found differences in the proportional representation of each clade when comparing Scotland with other UK regions in the period before the UoG residential halls cluster (Fig. 3).

Recent information from the Higher Education Statistics Agency estimates that around 84% of Scottish university students from within the UK are domiciled within Scotland (Supplementary Table S1, supplementary material). This compares to between 35 and 63% of students in England domiciled within the region of their higher education provider, 62% in Wales, and 93% in Northern Ireland. For UoG students specifically, 81% were domiciled in Scotland, with the percentage of students resident across the remaining UK regions ranging from 0.4 to 3% (see Supplementary Table S2, supplementary material). Employing a simple model taking into consideration the high percentage of UoG students domiciled within Scotland (Supplementary Tables S1–S2, supplementary material), we estimated a relatively high probability of Scottish relative to English origin for clades 1 and 3 (90:10 and 99:1 respectively), although these clades were not notably prevalent in Scotland in the

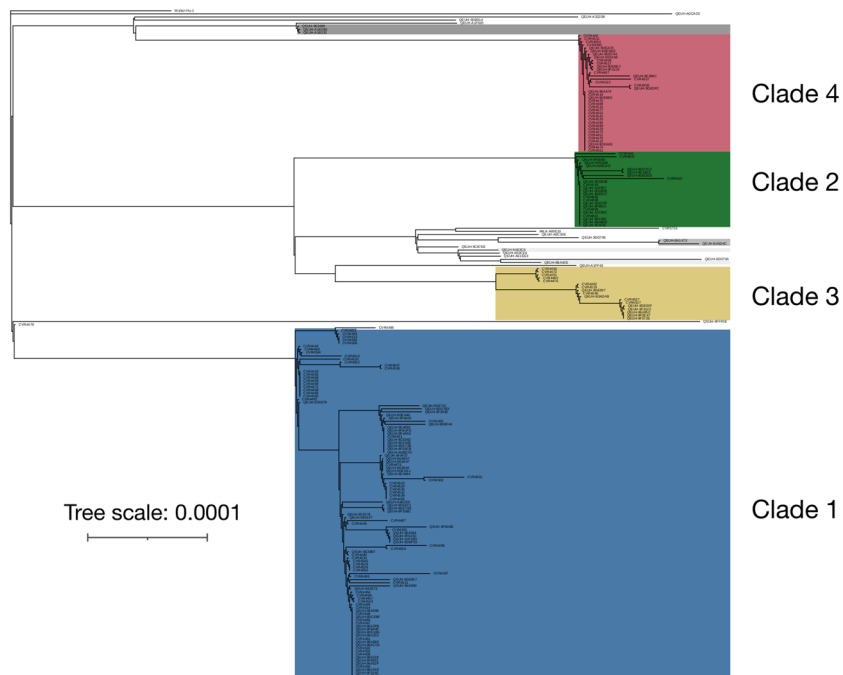


Figure 2. Phylogenetic tree of high quality genome sequences collected from University of Glasgow students. The reference sequence Wuhan Hu-1 (GenBank accession MN90894) was included as an outgroup. Four notable clades of sequences were identified, corresponding to cases that were putatively linked via direct transmission, as identified by the A2B-COVID software package. Shades of gray indicate three other sets of putatively linked cases. The phylogenetic analysis was conducted using iQ-Tree²⁸, while the figure was made with iTOL⁴⁴. The tree scale indicates the fraction of sites in the genome at which substitutions were observed. The tree scale bar denotes nucleotide substitutions per site.

period before the residential halls cluster (Fig. 3). Evidence was inconclusive for clades 2 and 4 (probability ratio of 60:40 in both cases) (see “Methods”).

Sequences from five residential halls (43 cases), and private accommodation (5 cases), were genetically indistinguishable from sequences identified from outbreaks associated with six other Scottish universities (Supplementary Fig. S4, supplementary material). These data may point towards acquisition of infection in Scottish domicile locations prior to the new academic term, as is the apparent case for at least two of the four clades.

The small number of UoG-outbreak-associated clades detected following the outbreak is consistent with this group of student cases having no long-term impact on the persistence of SARS-CoV-2 in the Scottish community.

Genetic diversity across halls suggests a high degree of student mixing. Sequences from each of the four major transmission clusters were found distributed across multiple residential halls and among students resident in private accommodation (Fig. 4). This picture is consistent with student-to-student transmission, either within residential halls and/or alongside common sources of exposure. The observation of mixed viral populations within most residential halls is consistent with each having had multiple SARS-CoV-2 introductions. Observations of viruses from different transmission clusters were seen in rapid succession within halls, making it difficult to identify the location of the first case in each and suggesting extensive mixing between students prior to the outbreaks. The rapid drop in cases suggests that transmission within these halls was quickly contained.

Cases infected with SARS-CoV-2 other than the four major clades were rarely associated with residential halls, with the vast majority resident in private accommodation (Fig. 4).

Evidence suggests a short-lived impact on local community case incidence. An analysis of student and community sequences showed considerable intermixing of cases, with multiple subclades of a constructed phylogenetic tree spanning both populations (Fig. 5). Sequences used for this analysis comprised the 112 student sequences from transmission cluster 1, and additionally the 104 sequences in the associated UK-wide phylogenetic clade that originated within NHS GGC. While our data are insufficient to derive accurate rates or directionality of transmission, the data are consistent with ongoing mixing of students with others in the local community, though with no long-term impact in terms of persisting chains of transmission.

Age-stratified proportions of SARS-CoV-2 infections identified among PCR-tested individuals in the City of Glasgow (excluding UoG students) showed a rise in cases among adults aged 17–24 years coinciding with the start of the UoG residential halls cluster (Fig. 6). Furthermore, the age distribution of sequenced community cases identified in NHS GGC falling into UK-wide clades associated with the four major student transmission

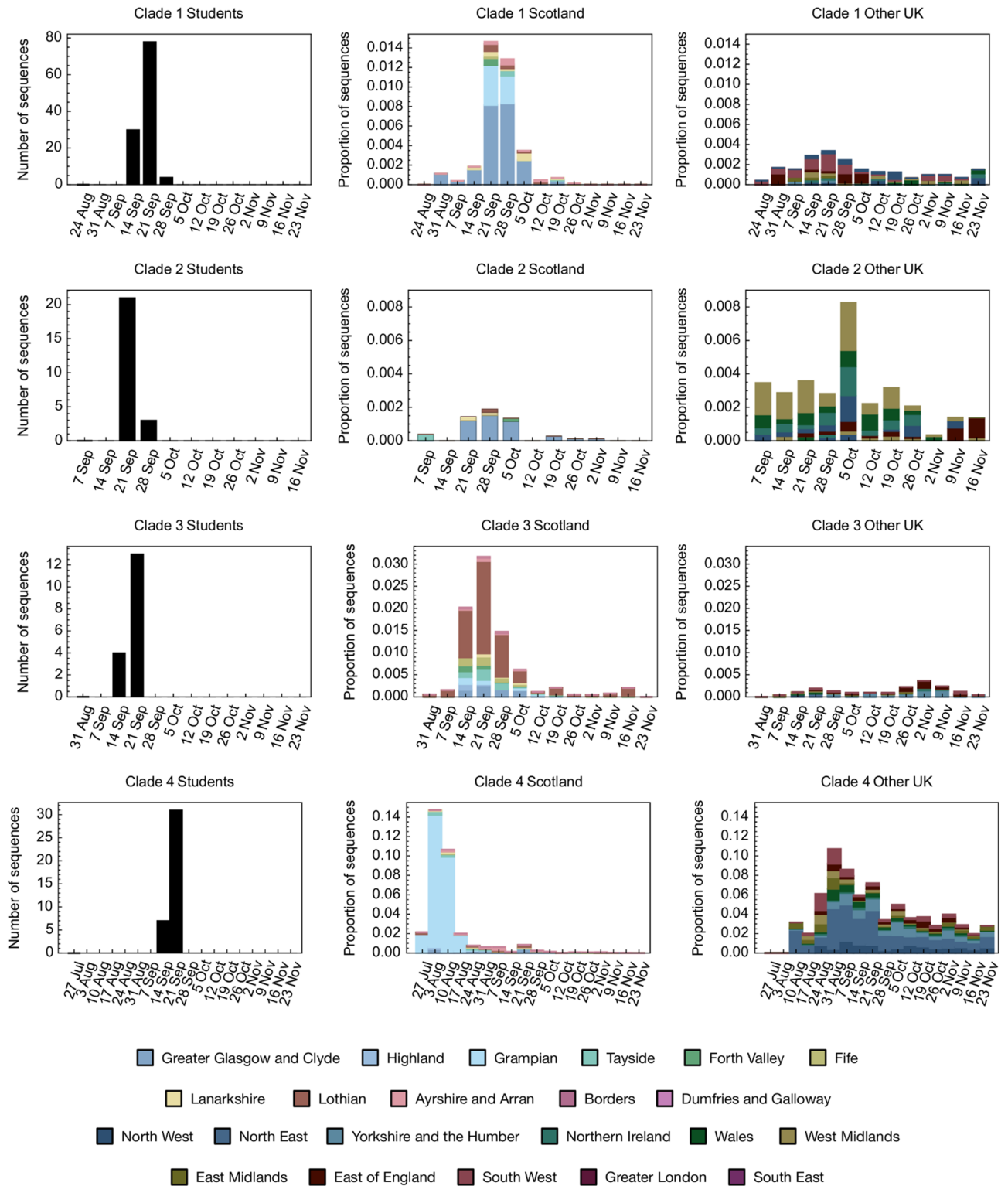


Figure 3. Potential origins of major SARS-CoV-2 clades identified in the University of Glasgow (UoG) student population. Numbers of UoG student sequences associated with each major clade (left column) are shown alongside data describing the associated UK-wide clades. Figures show the proportion of SARS-CoV-2 sequences collected each week in Scotland (middle column) and in the rest of the UK (right column) that were associated with each student clade, coloured by the regions in which the clade-associated sequences were collected. Total numbers of sequences informing this analysis were as follows: Clade 1 n = 412; Clade 2 n = 255; Clade 3 n = 541; Clade 4 n = 4006.

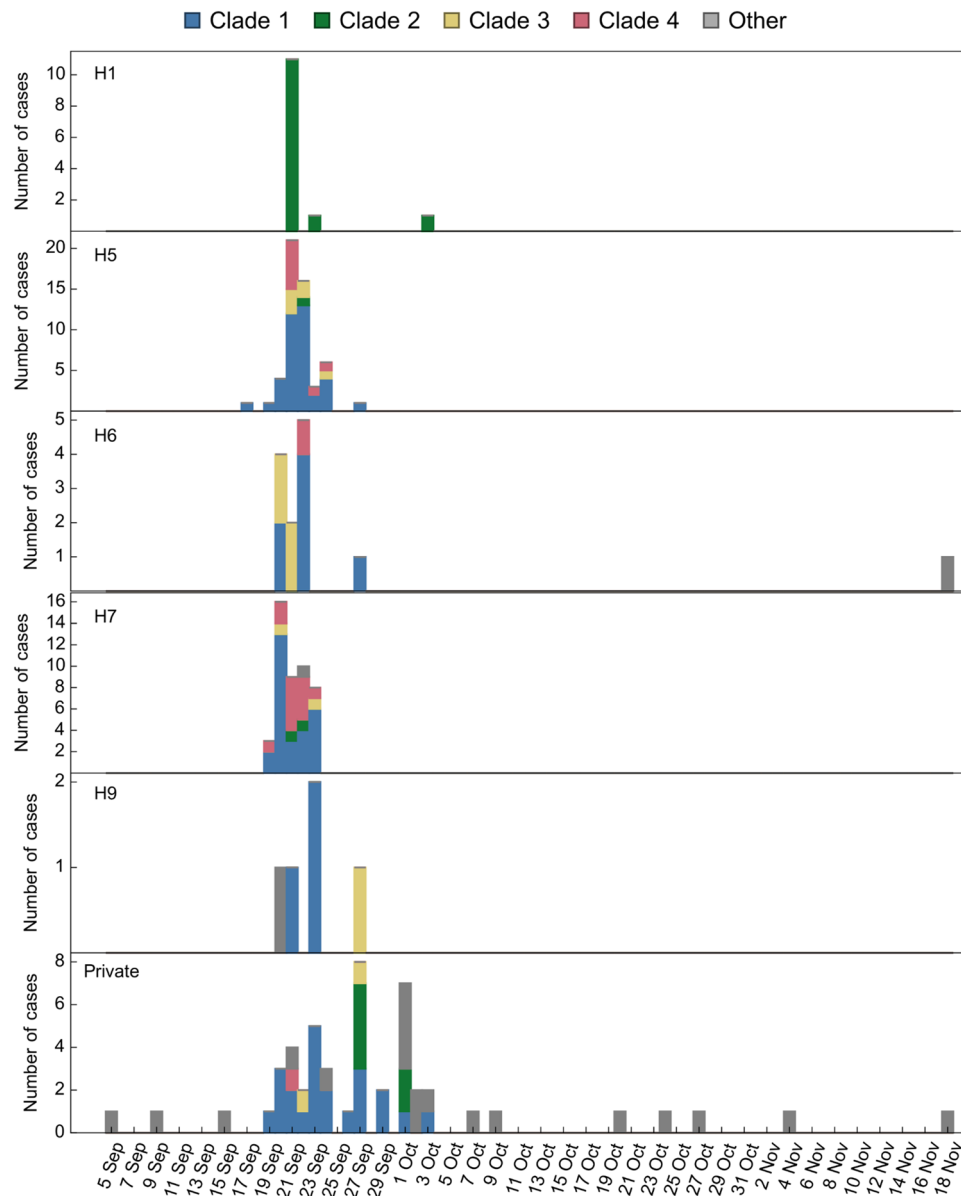


Figure 4. Dates of sequenced PCR-confirmed cases of SARS-CoV-2 infection identified in University of Glasgow students resident in halls or private accommodation, coloured by transmission clade. Data is shown for residential halls in which a minimum of four cases were observed. The identification of viruses from multiple transmission clusters in many of the residential halls suggests that there were multiple introductions of SARS-CoV-2 into each hall.

clusters showed a similar bias towards 17–24 year olds, consistent with mixing of students and the local community occurring predominantly among this age group (Fig. 7). The rise in the proportion of positive tests among 17–24 year olds was also seen on average across the rest of Scotland (Fig. 6), suggesting the burden of cases in this age group, with the potential impact of student populations at the start of term, was not unique to the City of Glasgow.

Discussion

We combined genome sequence data with epidemiological information to investigate a cluster of COVID-19 cases associated with UoG student residential halls. These cases occurred prior to rollout of the UK's COVID-19 vaccination programme and the use of lateral flow device tests for regular asymptomatic self-testing by the students. We evaluated frequency and origins of SARS-CoV-2 introductions into the student population, the likelihood of transmission among the students, and impact on the local community. The integration of genomic, evolutionary, and epidemiological approaches provides a deep insight into the nature of viral spread within a university accommodation context.

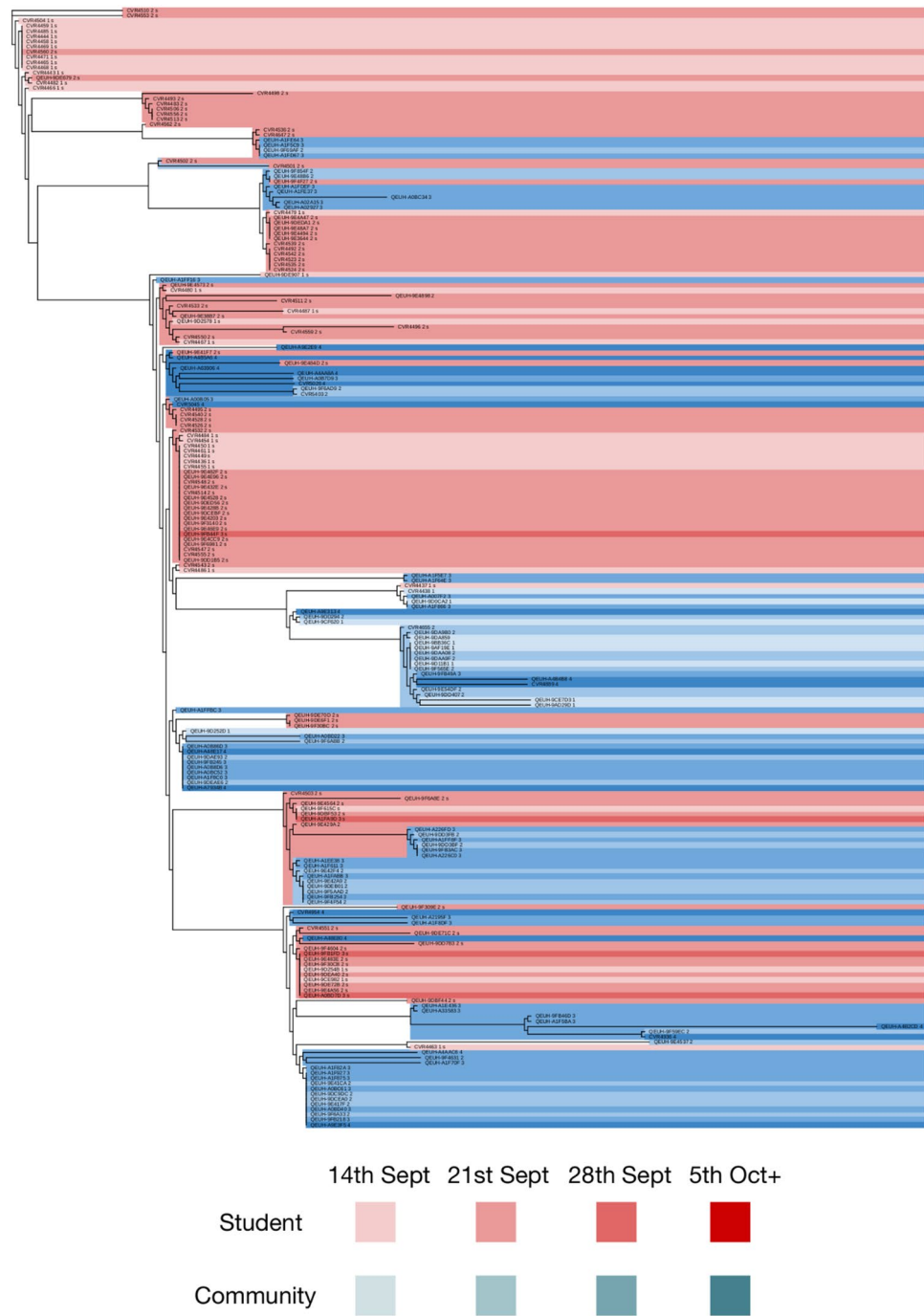


Figure 5. Phylogenetic relationship between student (red) and community (blue) sequences associated with transmission clade 1. Shades indicate the week in which sequences were collected, the lightest colour indicating dates before and including the week of 14th September, and the darkest colour indicating dates from the 5th October onwards. Community sequences describe those identified in NHS Greater Glasgow and Clyde sharing a common ancestor with the student sequences. Student and community cases were observed in multiple subclades of the tree, indicating ongoing transmission between the students and local community. Phylogenetic analysis was conducted using iQ-Tree³⁸, while the figure was made with iTOL⁴⁴.

Our study highlights several important features of public health significance. Firstly, the increase in SARS-CoV-2 infections among the student population primarily involved university accommodation. The sudden rise in student cases emerged shortly after social events linked to “Freshers’ week” at the start of the university term, providing a strong indication that the outbreaks spread via induction events and/or other social activity and gatherings²⁷.

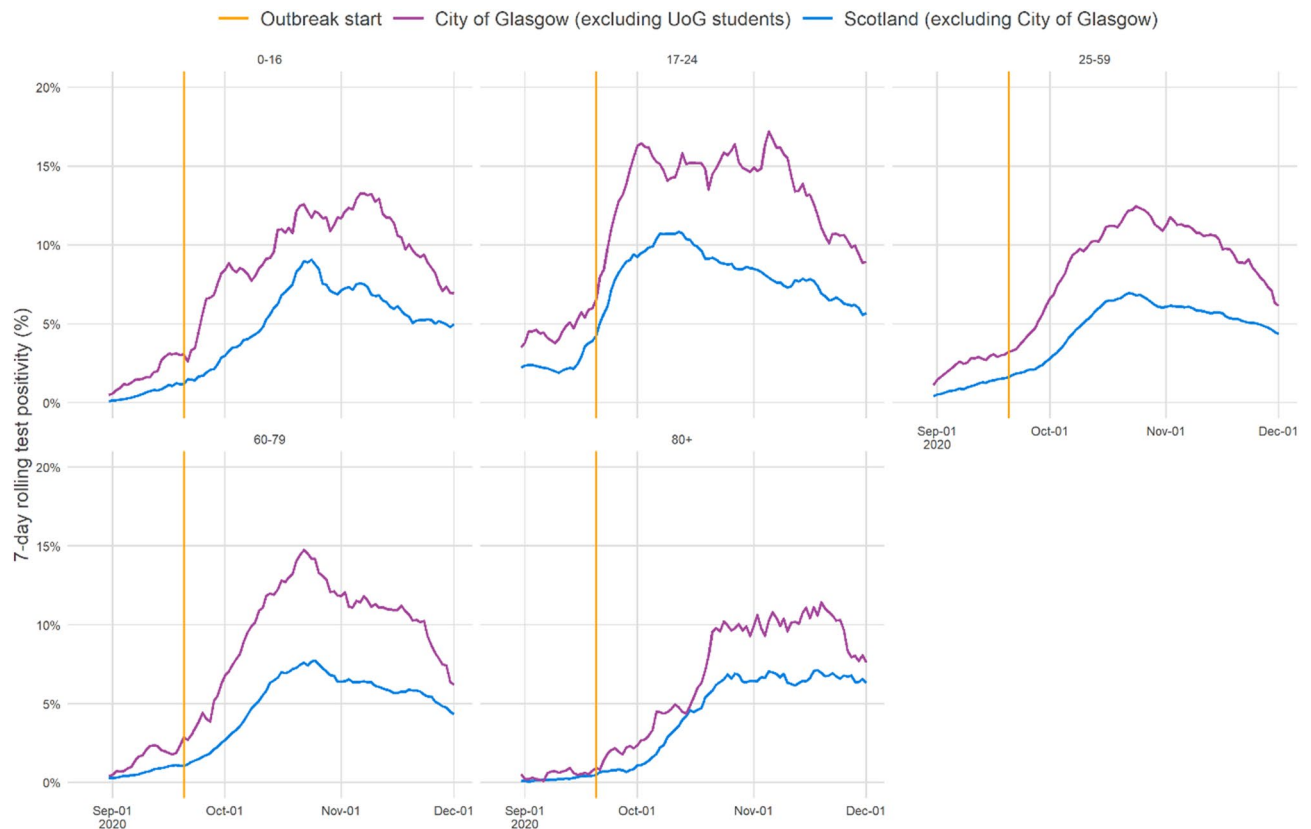


Figure 6. Age-stratified 7-day rolling proportions of PCR-confirmed cases of SARS-CoV-2 infection (test positivity) among all tested individuals in the City of Glasgow community [excluding University of Glasgow (UoG) students], and in comparison with the rest of Scotland, excluding the City of Glasgow. The vertical lines indicate the start of the UoG student halls cluster around 19th September 2020.

Secondly, while there was evidence of at least 11 introductions of SARS-CoV-2 into the university student population, our data were consistent with only four leading to substantial onward transmission among students, each consistent with a separate outbreak. Our data suggests the presence of large variations in the numbers of individuals infected by each introduction, potentially attributable to superspreading events whereby a small number of individuals cause a disproportionate number of cases²⁸. Although the sequencing of UoG student cases was incomplete, with around 30% coverage of the residential halls cluster, we can be reasonably confident that all major transmission clusters were detected.

Thirdly, the data were strongly indicative of introductions into the university from Scotland for at least two of the four clades. Our findings highlight that although the potential for virus importations into Scotland through long-distance student movements remains, the risk of new variants being introduced to Scotland at the start of the academic year would appear low in comparison to other UK regions given the high proportion of Scottish students domiciled within Scotland²⁹. The risk of importing SARS-CoV-2 from countries outside of the UK was not directly addressed by our study owing to a lack of robust data at a global level.

Fourthly, our study highlights the rapid spread of SARS-CoV-2 within and between university accommodation. Halls of residence provide household-like conditions enabling direct and indirect transmission of SARS-CoV-2 via respiratory (aerosol and droplet) and fomite routes. Accommodation at the UoG comprises predominantly off-campus accommodation, representing a mixture of self-contained flats, shared kitchen spaces and bathroom facilities, and common room/study areas²³. The identified transmission clusters were consistent with student–student transmission linked to household-like conditions and/or shared facilities within residential halls. Social gatherings may also be a conduit for the apparent transmission between households and likely contributed to the multiple viral populations within halls, with mixing of individuals returning to the university from various locations, and with compliance to social distancing expected to be poor at such events.

Despite the scale and speed at which the outbreak grew, the rapid curtailment of the outbreaks demonstrates that the swift action of the university and local Public Health Protection Unit, together with an apparent high degree of student compliance, prevented any further rise in case numbers and limited any impact on the local community. We note that a decline in proportional sequencing coverage in Scotland in the period subsequent to the UoG outbreaks (from an average daily coverage of ~24% in September 2020 to ~4% from mid-October 2020) limits the ability to detect associated variants circulating at low prevalence, and precludes a direct comparison of observed clade-specific case incidence rates between Scotland and England. However, we can be reasonably confident that variants associated with any significant outbreaks or long-lasting chains of transmission would be detectable among the community sequences available, given that sequencing was conducted daily over an

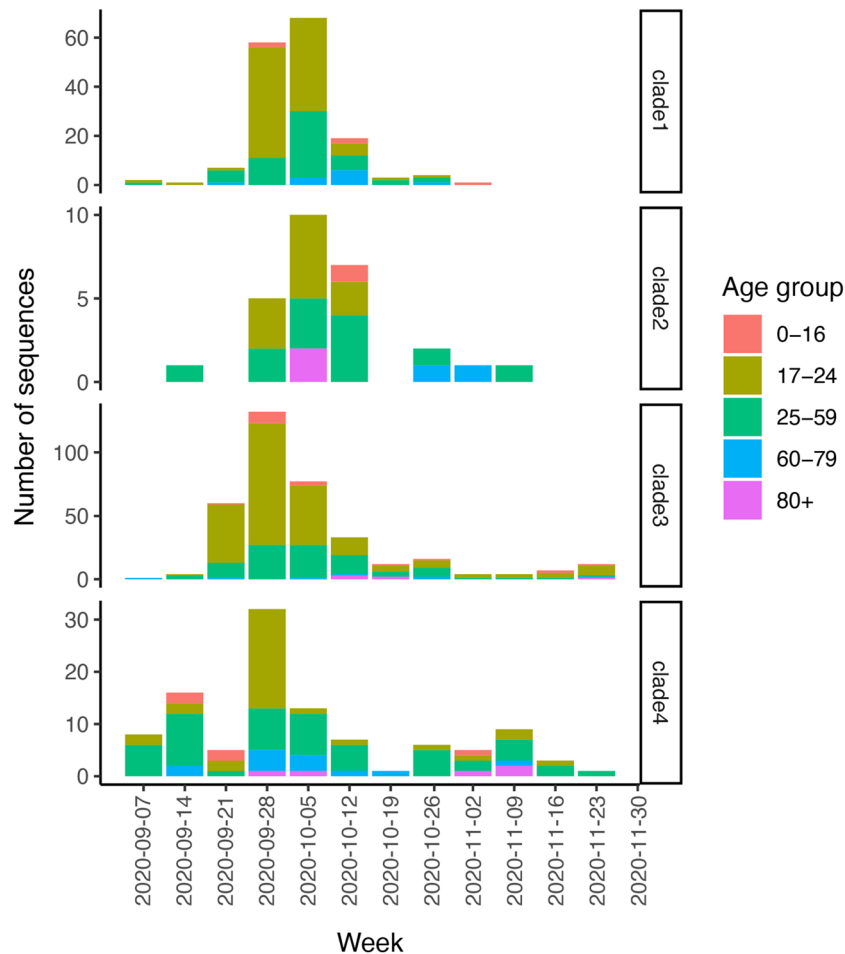


Figure 7. Age distribution of SARS-CoV-2 cases among residents of NHS Greater Glasgow and Clyde with sequences that phylogenetically clustered with each University of Glasgow (UoG) student transmission clade. UoG student cases were excluded. Numbers of sequences informing this analysis were as follows: Clade 1 n = 163, Clade 2 n = 27, Clade 3 n = 362, Clade 4 n = 106.

extended time frame. It is likely that self-isolation by students was the predominant factor leading to the rapid end of the student halls-associated outbreaks, although our study did not formally evaluate control measure effectiveness.

At the time of the UoG outbreaks, most educational institutes in the UK had undergone emergency closure of in-person teaching in response to the COVID-19 epidemic, resulting in prolonged periods of remote learning. However, significant outbreaks arose in the university setting, albeit during a period of relatively high susceptibility and prior to rollout of the national vaccination programme. In contrast to 2020/21, numbers of university student cases were comparatively lower in Scotland at the start of the 2021/22 academic year³⁰ in the absence of university closures and restriction of social activities. This period saw a decreasing case incidence nationally alongside the accumulation of immune protection via vaccination and naturally-acquired immunity, and with the introduction of self-testing by asymptomatic students using lateral flow device tests^{1,21,31}. The opening of educational settings during periods of low incidence may be an important component of risk mitigation³², alongside promotion of safer socialisation practices and enhanced social care and support structures³³.

The results from our study contrast with those reported for an outbreak among students at the University of Cambridge (UoC)¹⁵. UoC is an atypical UK higher educational institution in that in addition to education, accommodation and the social setting among students revolves around independent colleges. By contrast, UoG is far more centralised and integrated, albeit with university accommodation being more geographically intertwined with the local community. In Cambridge, the outbreak was dominated by a single large cluster that persisted until the national lockdown in November 2020. That the UoG residential halls were geographically restricted likely enabled greater control and enforcement of isolation measures through provision of student support. The outbreaks associated with UoG did however show greater interaction between students and the broader community than observed for UoC, with community and student cases in the largest clade being substantially mixed. The contrasting findings across studies provides invaluable context in understanding the role of SARS-CoV-2 spread within the university setting. Differences in university organisation and housing structure are likely to

have a pronounced impact on the potential spread of respiratory viruses, and the physical location of students within a city can affect both the nature and impact of an outbreak.

Our study has some important limitations. Firstly, a comparison of infection risks associated with student social activities, or the household conditions and/or shared facilities within residential halls, was not possible. Secondly, interpretation of direct student-to-student transmission events is limited by lack of information on how sequenced cases were linked to individual households within a residential hall, and the low levels of variation in SARS-CoV-2 genomes make it difficult to conclusively define epidemiological links. Thirdly, the analysis of the origins of the university student associated clades does not consider the potential for the introduction of cases into the student population via international travel. Fourthly, it is possible that some student cases were misclassified as community cases. Fifth, we note that the presence, frequency and diversity of SARS-CoV-2 may be biased by the targeted sampling of outbreaks, particularly during periods of low sequencing coverage. Some targeted sequencing of the UoG student outbreak had occurred to assist with initial investigations in near real-time. Finally, only 30% of identified student cases were sequenced. We have aimed to provide conservative estimates throughout.

These analyses have provided valuable insight into a large outbreak in an off-campus university accommodation setting, and an evidence base to inform future policy recommendations for students returning to universities at the start of terms. The higher education setting presents a risk for contributing to the winter burden of COVID-19. With rapid identification, and implementation of non-pharmaceutical control interventions, the impact of outbreaks on local communities may be limited. However, while high rates of vaccination have moderated the impact of SARS-CoV-2 disease severity on the UK population, it is clear SARS-CoV-2 will continue to evolve in the human population such that student populations, as with all demographics, should be encouraged to be up to date with vaccinations (currently three doses in under 75 year olds as of mid-2022).

Methods

Epidemiological investigation. A retrospective search was performed in a database of NHS Scotland Test and Protect contact tracing interviews, known as the Case Management System. The case definition pertaining to a University of Glasgow (UoG) associated student was applied as follows: the text “Glasgow” and “university or uni” under occupation related fields, and with RT-PCR diagnostic test confirmation of a SARS-CoV-2 infection between 1st September 2020 and 30th November 2020. Free text notes were inspected and cases were excluded if determined to not be a University of Glasgow student. Residential postcodes were then matched to UoG student halls postcodes obtained from the Scottish Government Advanced Learning and Skills Analysis (ALSA), identifying cases resident in university accommodation. Further cross checking was performed: student status was corroborated by the individual’s age (with the majority of student cases ≤ 19 years) and/or an additional interrogation of free text notes. Four cases had missing age information. We note that our approach is relatively conservative, and the number of student cases identified may differ from that based on other methods (such as reported in Ref.²⁰).

Student SARS-CoV-2 whole genome sequences. In Scotland, SARS-CoV-2 genome sequencing is coordinated by the NHS Sequencing Service. Respiratory specimens collected through routine RT-PCR NHS diagnostic testing and UK Government community testing were submitted to The COVID-19 Genomics UK Consortium (COG-UK) partnership laboratories at the MRC-University of Glasgow Centre for Virus Research and Sanger Institute for sequencing. The genomic sequence data were uploaded to CLIMB (Cloud Infrastructure for Microbial Bioinformatics)-COVID³⁴ and amalgamated with all Scottish sequences. The UoG student case list was linked to Scottish genomic sequence data derived from respiratory specimens collected between 1st September 2020 and 30th November 2020 using a unique identifier. Out of a total 220 identified sequences, 214 were considered high quality for further analysis, with $>80\%$ nucleotide coverage across the viral genome. Four sequenced cases had specimen dates preceding 19th September.

Inference of transmission clusters among UoG students. Putative transmission clusters were inferred using the A2B-COVID software package¹⁷. Given symptom onset dates and SARS-CoV-2 genome sequences from multiple individuals, this carries out a series of hypothesis tests, calculating whether the data from each pair of individuals A and B is statistically consistent with a model of A having directly infected B; data are assessed as being consistent with transmission, borderline, or unlikely to have arisen from a direct transmission event. Our model utilises distributions for the infectivity profile and time to symptom onset described for SARS-CoV-2 by other publications^{35–37}. In the absence of symptom onset dates, respiratory specimen collection dates were used as proxy estimates. A sub-setting procedure was used to generate putative transmission clusters, placing two individuals in the same cluster if transmission between those two individuals in either direction was not unlikely.

Phylogenetic inference of UoG student sequences was then conducted using the IQ-Tree2 package³⁸ in order to characterise the transmission clusters in terms of phylogenetic clades. Inferences used ModelFinder to identify the most appropriate model in each case³⁹. The trees of Figs. 2 and 5 were thus inferred using a TIM2 model with empirical base frequencies and allowing for a proportion of invariable sites.

Inference of transmission clade origins. The sampling locations of sequences in each clade were used to determine the likely origin of each transmission cluster. Viral genome sequence data were then used to estimate the probability that each of the four major clades we identified arose from an introduction into the student population either from England or Scotland. To derive an estimate of the probability that the introduction of a clade

C came from England, we first estimated the fraction of genome sequences of clade C in England and Scotland in the two weeks prior to the observation of the first student case in that clade:

$$Q_{S,C} = \frac{g_{S,C}}{G_S},$$

and

$$Q_{E,C} = \frac{g_{E,C}}{G_E},$$

where $g_{S,C}$ and $g_{E,C}$ were the number of virus genome sequences in clade C collected in Scotland and England respectively, during the two weeks prior to the first observation of the clade in students at the UoG, while G_S and G_E were the total number of virus genome sequences collected in Scotland and England during the same period.

Using these values, we estimated the prevalence of each clade in Scotland and England to inform the probabilities $P_{S,C}$ and $P_{E,C}$ that a given individual in either Scotland or England respectively was infected with SARS-CoV-2 from clade C:

$$P_{S,C} = \frac{Q_{S,C}I_S}{N_S},$$

and

$$P_{E,C} = \frac{Q_{E,C}I_E}{N_E},$$

where I_S and I_E were the mean total numbers of positive tests reported per week in Scotland and England during the 2 weeks in question, as reported in UK Government statistics⁴⁰, while N_S and N_E are ONS mid-2020 population estimates for Scotland and England⁴¹. Given these values, we estimated the probability that a student arriving in Glasgow was infected with SARS-CoV-2 of clade C and domiciled in either Scotland ($X_{S,C}$) or England ($X_{E,C}$). For $X_{S,C}$ this is given by

$$X_{S,C} = \frac{P_{S,C}P_{S,U}}{P_{S,C}P_{S,U} + P_{E,C}P_{E,U}},$$

where $P_{S,U}$ and $P_{E,U}$ are the probabilities that a given student at UoG was domiciled in Scotland or England respectively. The equivalent value $X_{E,C}$ that the student was domiciled in England is given by $X_{E,C} = 1 - X_{S,C}$.

In our calculation we neglected the possibility of introductions from outside of either England or Scotland due to the lack of data from other locations. Our measure is approximate, being prone to sampling bias and limited data, but was used to provide a broad indication of the geographical origins of clades.

Phylogenetic inference for NHSGGC sequences. The COG-UK phylogeny for SARS-CoV-2 was produced using CLIMB-COVID data available on the 2020–12–31. The grapevine pipeline⁴² was then used for extracting the smallest phylogenetic clade containing the sequences in each putative transmission cluster. This was used to put the SARS-CoV-2 genome sequences from UoG students in the broader context of UK wide sequences. The clades were pruned to include only sequences from UoG students and NHSGGC and students from other Scottish Universities determined through contact tracing. For visualisation of these clades, identical sequences were collapsed and the figure was produced in The Environment for Tree Exploration⁴³.

Local epidemiological context in the City of Glasgow. All individuals with respiratory specimens collected for diagnostic testing of SARS-CoV-2 infections by RT-PCR in Scotland between 1st September 2020 and 30th November 2020 were identified from the national Electronic Communication of Surveillance in Scotland (ECOSS) database. Age-stratified 7-day rolling proportions of SARS-CoV-2 positive cases, among all RT-PCR-tested individuals, were estimated for the City of Glasgow, excluding cases with residential postcodes associated with UoG residential halls, in a conservative approach to removing the influence of UoG students from the analysis. A comparison was made with the rest of Scotland, excluding individuals resident in the City of Glasgow. The 17–24 years group represent the age range of the UoG residential halls-associated cases.

Ethics declaration. This study was undertaken as part of public health surveillance activity within the COVID-19 programme of Public Health Scotland, in line with the necessary associated regulations and guidelines. The retention and processing of information on individuals is conducted by Public Health Scotland as part of COVID-19 surveillance in Scotland in the context of emergency data processing (<https://www.informationgovernance.scot.nhs.uk/covid-19-privacy-statement/>), including the Civil Contingencies Act 2004, the NHS (Scotland) Act 1978 and the Public Health (Scotland) Act 2008, and under Articles 6(1)(e), 9(2)(h), 9(2)(i), 9(2)(j) of the General Data Protection Regulation. Surveillance data was shared with NHS Scotland according to the Intra NHS Scotland Data Sharing Accord (<https://www.informationgovernance.scot.nhs.uk/wp-content/uploads/2020/06/2020-06-17-Intra-NHS-Scotland-Sharing-Accord-v2.0.pdf>). Ethics approval was not required for this work which was based on pre-existing routine surveillance data for the Scottish population.

Data availability

The SARS-CoV-2 genome sequences used in this study are available from GISAID with identifiers provided in supplementary material. Other healthcare data are available upon application to the NHS Scotland Public Benefit and Privacy Panel For Health and Social Care: <https://www.informationgovernance.scot.nhs.uk/pbpphsc/>. The A2B-COVID tool is available from <https://chjackson.github.io/a2bcovid/>.

Received: 14 March 2022; Accepted: 27 June 2022

Published online: 19 July 2022

References

1. COVID-19 Daily Dashboard. COVID-19 vaccination in Scotland. Daily update. https://public.tableau.com/app/profile/phs.covid.19/viz/COVID-19DailyDashboard_15960160643010/Overview#!/vizhome/COVID-19DailyDashboard_15960160643010/Overview. (accessed 07 June 2022).
2. Coronavirus (COVID-19) Infection Survey, antibody and vaccination data, UK—Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveyantibodyandvaccinationdatafortheuk/4august2021>. (accessed 07 June 2022).
3. Antonelli, M. *et al.* Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: A prospective, community-based, nested, case-control study. *Lancet Infect. Dis.* **22**, 43–55 (2022).
4. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).
5. Thomson, E. C. *et al.* Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171–1187.e20 (2021).
6. Donaldson, A. L., Harris, J. P., Vivancos, R. & O'Brien, S. J. Risk factors associated with outbreaks of seasonal infectious disease in school settings, England, UK. *Epidemiol. Infect.* **148**, e287 (2020).
7. Avoid non-essential travel to Blackpool and other high risk areas—News—Public Health Scotland. <https://publichealthscotland.scot/news/2020/october/avoid-non-essential-travel-to-blackpool-and-other-high-risk-areas/>.
8. Marsh, K., Griffiths, E., Young, J. J., Gibb, C.-A. & McMenamin, J. Contributions of the EURO 2020 football championship events to a third wave of SARS-CoV-2 in Scotland, 11 June to 7 July 2021. *Euro Surveill.* **26**, 2100707 (2021).
9. COVID-19 clusters and outbreaks in occupational settings in the EU/EEA and the UK. <https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-in-occupational-settings.pdf>. (accessed 07 June 2022).
10. Pokhrel, S. & Chhetri, R. A literature review on impact of COVID-19 pandemic on teaching and learning. *High. Educ. Future* **8**, 133–141 (2021).
11. Principles for Managing SARS-CoV-2 Transmission Associated with Higher Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/914978/S0728_Principles_for_Managing_SARS-CoV-2_Transmission_Associated_with_Higher_Education_.pdf. (accessed 07 June 2022).
12. Hill, E. M., Atkins, B. D., Keeling, M. J., Tildesley, M. J. & Dyson, L. Modelling SARS-CoV-2 transmission in a UK university setting. *Epidemics* **36**, 100476 (2021).
13. Enright, J. *et al.* SARS-CoV-2 infection in UK university students: Lessons from September–December 2020 and modelling insights for future student return. *R. Soc. Open Sci.* **8**, 210310 (2021).
14. Beale, S., Hayward, A., Shallcross, L., Aldridge, R. W. & Fragaszy, E. A rapid review and meta-analysis of the asymptomatic proportion of PCR-confirmed SARS-CoV-2 infections in community settings. *Wellcome Open Res.* **5**, 266 (2020).
15. Aggarwal, D. *et al.* Genomic epidemiology of SARS-CoV-2 in a UK university identifies dynamics of transmission. *Nat. Commun.* **13**, 751 (2022).
16. Vusirikala, A. *et al.* Seroprevalence of SARS-CoV-2 antibodies in university students: Cross-sectional Study, December 2020, England. SSRN J. <https://doi.org/10.2139/ssrn.3787684> (2021).
17. Illingworth, C. J. R. *et al.* A2B-COVID: A method for evaluating potential SARS-CoV-2 transmission events. *medRxiv* <https://doi.org/10.1101/2020.10.26.20219642> (2020).
18. Meredith, L. W. *et al.* Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: A prospective genomic surveillance study. *Lancet Infect. Dis.* **20**, 1263–1272 (2020).
19. Aggarwal, D. *et al.* The role of viral genomics in understanding COVID-19 outbreaks in long-term care facilities. *Lancet Microbe* [https://doi.org/10.1016/S2666-5247\(21\)00208-1](https://doi.org/10.1016/S2666-5247(21)00208-1) (2021).
20. Public Health Scotland. COVID-19 in university accommodation settings. <https://www.publichealthscotland.scot/media/2905/covid-19-in-university-accommodation-settings-dec20-english.pdf>. (accessed 07 June 2022).
21. Public Health Scotland. COVID-19 Daily Dashboard. COVID-19 in Scotland. Trends and demographics. https://public.tableau.com/app/profile/phs.covid.19/viz/COVID-19DailyDashboard_15960160643010/Overview. (accessed 07 June 2022).
22. University of Glasgow—MyGlasgow—Planning, Insight & Analytics—Headcount Student Data (QlikView)—Domicile—Headcount of New Undergraduate Students by Domicile Breakdown 2011–12 to 2020–21. <https://www.gla.ac.uk/myglasgow/planning/qv/dom/d3/>. (accessed 07 June 2022).
23. University of Glasgow—MyGlasgow—Accommodation Services. <https://www.gla.ac.uk/myglasgow/accommodation/>. (accessed 07 June 2022).
24. University of Glasgow—MyGlasgow—Planning, Insight & Analytics—Headcount Student Data (QlikView). <https://www.gla.ac.uk/myglasgow/planning/qv/>. (accessed 07 June 2022).
25. Cov-Lineages. https://cov-lineages.org/lineage_list.html. (accessed 07 June 2022).
26. Rambaut, A. *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
27. Coronavirus: Several “parties” at Glasgow student halls broken up by police|HeraldScotland. <https://www.heraldscotland.com/news/18751219.coronavirus-several-parties-glasgow-student-halls-broken-police/>. (accessed 07 June 2022).
28. Adam, D. C. *et al.* Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat. Med.* **26**, 1714–1719 (2020).
29. Table 11—HE student enrolments by domicile and region of HE provider 2014/15 to 2019/20|HESA. <https://www.hesa.ac.uk/data-and-analysis/students/table-11>. (accessed 07 June 2022).
30. Public Health Scotland COVID-19 Statistical Report. <https://publichealthscotland.scot/media/9908/21-10-27-covid19-publication-report.pdf>. (accessed 07 June 2022).
31. University of Glasgow—MyGlasgow—Coronavirus: information for students and staff—Booking a test. <https://www.gla.ac.uk/myglasgow/coronavirus/testing/>. (accessed 07 June 2022).
32. Yamey, G. & Walensky, R. P. Covid-19: Re-opening universities is high risk. *BMJ* **370**, m3365 (2020).
33. The Independent SAGE Report 9. <https://www.independentsage.org/wp-content/uploads/2020/09/Independent-SAGE-Universities-Consultation.pdf>. (accessed 07 June 2022).

34. Nicholls, S. M. *et al.* CLIMB-COVID: Continuous integration supporting decentralised sequencing for SARS-CoV-2 genomic surveillance. *Genome Biol.* **22**, 196 (2021).
35. Ashcroft, P. *et al.* COVID-19 infectivity profile correction. *Swiss. Med. Wkly.* **150**, w20336 (2020).
36. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
37. Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
38. Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
39. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
40. Cases in the UK | Coronavirus in the UK. <https://coronavirus.data.gov.uk/details/cases>. (accessed 07 June 2022).
41. Population estimates for the UK, England and Wales, Scotland and Northern Ireland - Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annu>. (accessed 07 June 2022).
42. GitHub - COG-UK/grapevine: Process consensus file for COG-UK data. <https://github.com/COG-UK/grapevine>. (accessed 07 June 2022)
43. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638 (2016).
44. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

Acknowledgements

We thank the Public Health Microbiology and Real-Time Epidemiology teams at Public Health Scotland, and the Public Health Protection Unit at NHS Greater Glasgow and Clyde, who developed and managed the surveillance infrastructure and data sharing that enabled this study as part of the COVID-19 response in Scotland. We also thank the University of Glasgow, in particular Selina Woolcott and Maureen Hood, for providing helpful contextual information. The CVR authors are supported by the Medical Research Council [MC_UU_12014/12 and MC_UU_12018/12]. COG-UK is supported by funding from the Medical Research Council (MRC) part of UK Research & Innovation (UKRI), the National Institute of Health Research (NIHR) [Grant code: MC_PC_19027], and Genome Research Limited, operating as the Wellcome Sanger Institute.

Author contributions

S.N., C.M., E.T., D.L.R., M.T.G.H. and A.S.P. conceived the study. S.N., J.H., C.I., A.S.P. designed the analysis. S.N., E.G., A.S.P. performed epidemiological investigation. K.S., K.N., N.P., L.T., S.C., V.B.S., R.O., E.J.G., R.M.T. performed generation of student genome sequence data. S.N., J.H., S.S. performed data cleaning and curation. S.N., N.C., J.E., J.H., C.I. performed analyses. K.T., R.N.G., A.S.F., E.T., D.L.R., M.T.G.H. coordinated and supervised data generation. S.N., C.I., A.S.P. supervised analyses. S.N., C.I. wrote the first draft. All authors interpreted data and contributed to and approved the final version. The COVID-19 Genomics UK Consortium (COG-UK) coordinated the collection and generation of the UK-wide genome sequences used in this study.

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/s41598-022-15661-1>.

Correspondence and requests for materials should be addressed to S.N. or C.J.R.I.

Reprints and permissions information is available at 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

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

Samuel C. Robson^{19,90}, Thomas R. Connor^{58,80}, Nicholas J. Loman⁴⁹, Tanya Golubchik¹²³, Rocio T. Martinez Nunez⁵², David Bonsall¹²³, Andrew Rambaut¹¹⁰, Luke B. Snell¹⁸, Rich Livett¹²², Catherine Ludden^{26,76}, Sally Corden⁸⁰, Eleni Nastouli^{36,101,102}, Gaia Nebbia¹⁸, Ian Johnston¹²², Katrina Lythgoe¹²³, M. Estee Torok^{25,26}, Ian G. Goodfellow³⁰, Jacqui A. Prieto^{88,103}, Kordo

Saeed^{89,103}, David K. Jackson¹²², Catherine Houlihan^{101,102}, Dan Frampton^{100,101}, William L. Hamilton²⁵, Adam A. Witney⁴⁷, Giselda Bucca¹⁰⁷, Cassie F. Pope^{46,47}, Catherine Moore⁸⁰, Emma C. Thomson³, Ewan M. Harrison^{108,122}, Colin P. Smith¹⁰⁷, Fiona Rogan⁸³, Shaun M. Beckwith¹⁴, Abigail Murray¹⁴, Dawn Singleton¹⁴, Kirstine Eastick⁴³, Liz A. Sheridan¹⁰⁴, Paul Randell¹⁰⁵, Leigh M. Jackson¹¹¹, Cristina V. Ariani¹²², Sónia Gonçalves¹²², Derek J. Fairley^{59,83}, Matthew W. Loose²⁴, Joanne Watkins⁸⁰, Samuel Moses^{31,112}, Sam Nicholls⁴⁹, Matthew Bull⁸⁰, Roberto Amato¹²², Darren L. Smith^{42,71,72}, David M. Aanensen^{20,122}, Jeffrey C. Barrett¹²², Dinesh Aggarwal^{126,122,76}, James G. Shepherd³, Martin D. Curran⁷⁷, Surendra Parmar⁷⁷, Matthew D. Parker¹¹⁵, Catryn Williams⁸⁰, Sharon Glaysher⁷⁴, Anthony P. Underwood^{20,122}, Matthew Bashton^{42,71}, Nicole Pacchiarini⁸⁰, Katie F. Loveson⁹⁰, Matthew Byott^{101,102}, Alessandro M. Carabelli²⁶, Kate E. Templeton^{62,110}, Thushan I. de Silva¹¹⁵, Dennis Wang¹¹⁵, Cordelia F. Langford¹²², John Sillitoe¹²², Rory N. Gunson⁶¹, Simon Cottrell⁸⁰, Justin O'Grady^{81,109}, Dominic Kwiatkowski^{114,122}, Patrick J. Lillie⁴³, Nicholas Cortes³⁹, Nathan Moore³⁹, Claire Thomas³⁹, Phillipa J. Burns⁴³, Tabitha W. Mahungu⁸⁶, Steven Liggett⁹², Angela H. Beckett^{19,87}, Matthew T. G. Holden¹, Lisa J. Levett⁴⁰, Husam Osman^{41,76}, Mohammed O. Hassan-Ibrahim¹⁰⁵, David A. Simpson⁸³, Mrera Chand⁷⁸, Ravi K. Gupta¹⁰⁸, Alistair C. Darby¹¹³, Steve Paterson¹¹³, Oliver G. Pybus²⁹, Erik M. Volz⁴⁵, Daniela de Angelis¹¹, David L. Robertson³, Andrew J. Page⁸¹, Inigo Martincorena¹²², Louise Aigrain¹²², Andrew R. Bassett¹²², Nick Wong⁵⁶, Yusri Taha⁹⁵, Michelle J. Erkiert¹⁰⁵, Michael H. Spencer Chapman^{108,122}, Rebecca Dewar⁶², Martin P. McHugh^{62,117}, Siddharth Mookerjee^{44,63}, Stephen Aplin¹⁰³, Matthew Harvey¹⁰³, Thea Sass¹⁰³, Helen Umpleby¹⁰³, Helen Wheeler¹⁰³, James P. McKenna⁵⁹, Ben Warne⁷⁹, Joshua F. Taylor²⁸, Yasmin Chaudhry³⁰, Rhys Izuagbe³⁰, Aminu S. Jahun³⁰, Gregory R. Young^{42,71}, Claire McMurray⁴⁹, Clare M. McCann^{71,72}, Andrew Nelson^{71,72}, Scott Elliott⁷⁴, Hannah Lowe³¹, Anna Price⁵⁸, Matthew R. Crown⁷¹, Sara Rey⁸⁰, Sunando Roy¹⁰², Ben Temperton¹¹¹, Sharif Shaaban¹, Andrew R. Hesketh¹⁰⁷, Kenneth G. Laing⁴⁷, Irene M. Monahan⁴⁷, Judith Heaney^{40,101,102}, Emanuela Pelosi¹⁰³, Siona Silveira¹⁰³, Eleri Wilson-Davies¹⁰³, Helen Fryer¹²³, Helen Adams¹³, Louis du Plessis²⁹, Rob Johnson⁴⁵, William T. Harvey^{2,3}, Joseph Hughes^{3,125}, Richard J. Orton³, Lewis G. Spurgin⁶⁵, Yann Bourgeois⁸⁷, Chris Ruis¹⁰⁸, Áine O'Toole¹¹⁰, Marina Gourtovaia¹²², Theo Sanderson¹²², Christophe Fraser¹²³, Jonathan Edgeworth¹⁸, Judith Breuer^{35,102}, Stephen L. Michell¹¹¹, John A. Todd¹²¹, Michaela John¹⁷, David Buck¹²¹, Kavitha Gajee⁴³, Gemma L. Kay⁸¹, Sharon J. Peacock^{26,76}, David Heyburn⁸⁰, Katie Kitchman⁴³, Alan McNally^{49,99}, David T. Pritchard⁵⁶, Samir Dervisevic⁶⁴, Peter Muir⁷⁶, Esther Robinson^{41,76}, Barry B. Vipond⁷⁶, Newara A. Ramadan⁸⁴, Christopher Jeanes⁹⁶, Danni Weldon¹²², Jana Catalan¹²⁴, Neil Jones¹²⁴, Ana da Silva Filipe³, Chris Williams⁸⁰, Marc Fuchs⁸³, Julia Miskelly⁸³, Aaron R. Jeffries¹¹¹, Karen Oliver¹²², Naomi R. Park¹²², Amy Ash¹², Cherian Koshy¹², Magdalena Barrow¹⁵, Sarah L. Buchan¹⁵, Anna Mantzouratou¹⁵, Gemma Clark²¹, Christopher W. Holmes²², Sharon Campbell²³, Thomas Davis²⁷, Ngee Keong Tan²⁸, Julianne R. Brown³⁵, Kathryn A. Harris^{35,48}, Stephen P. Kidd³⁹, Paul R. Grant⁴⁰, Li Xu-McCrae⁴¹, Alison Cox^{44,69}, Pinglawathee Madona^{44,69}, Marcus Pond^{44,69}, Paul A. Randell^{44,69}, Karen T. Withell⁵⁴, Cheryl Williams⁵⁷, Clive Graham⁶⁶, Rebecca Denton-Smith⁶⁸, Emma Swindells⁶⁸, Robyn Turnbull⁶⁸, Tim J. Sloan⁷³, Andrew Bosworth^{41,76}, Stephanie Hutchings⁷⁶, Hannah M. Pymont⁷⁶, Anna Casey⁸², Liz Ratcliffe⁸², Christopher R. Jones^{85,111}, Bridget A. Knight^{85,111}, Tanzina Haque⁸⁶, Jennifer Hart⁸⁶, Dianne Irish-Tavares⁸⁶, Eric Witele⁸⁶, Craig Mower⁹², Louisa K. Watson⁹², Jennifer Collins⁹⁵, Gary Eltringham⁹⁵, Dorian Crudgington¹⁰⁴, Ben Macklin¹⁰⁴, Miren Iturriza-Gomara¹¹³, Anita O. Lucaci¹¹³, Patrick C. McClure¹¹⁹, Matthew Carlile²⁴, Nadine Holmes²⁴, Christopher Moore²⁴, Nathaniel Storey³⁵, Stefan Rooke¹, Gonzalo Yebra¹, Noel Craine⁸⁰, Malorie Perry⁸⁰, Nabil-Fareed Alikhan⁸¹, Stephen Bridgett⁸³, Kate F. Cook⁹⁰, Christopher Fearn⁹⁰, Salman Goudarzi⁹⁰, Ronan A. Lyons⁹⁴, Thomas Williams¹¹⁰, Sam T. Haldenby¹¹³, Jillian Durham¹²², Steven Leonard¹²², Robert M. Davies¹²², Rahul Batra¹⁸, Beth Blane²⁶, Moira J. Spyer^{36,101,102}, Perminder Smith^{38,118}, Mehmet Yavus^{91,115}, Rachel J. Williams¹⁰², Adhyana I. K. Mahanama¹⁰³, Buddhini Samaraweera¹⁰³, Sophia T. Girgis¹⁰⁸, Samantha E. Hansford¹¹⁵, Angie Green¹²¹, Charlotte Beaver¹²², Katherine L. Bellis^{108,122}, Matthew J. Dorman¹²², Sally Kay¹²², Liam Prestwood¹²², Shavanthi Rajatileka¹²², Joshua Quick⁴⁹, Radoslaw Poplawski⁴⁹, Nicola Reynolds¹⁶, Andrew Mack⁵⁸, Arthur Morriss⁵⁸, Thomas Whalley⁵⁸, Bindi Patel¹⁸, Iliana Georgana³⁰, Myra Hosmillo³⁰, Malte L. Pinckert³⁰, Joanne Stockton⁴⁹, John H. Henderson⁷¹, Amy Hollis⁷¹, William Stanley⁷¹, Wen C. Yew⁷¹, Richard

Myers⁷⁸, Alicia Thornton⁷⁸, Alexander Adams⁸⁰, Tara Annett⁸⁰, Hibo Asad⁸⁰, Alec Birchley⁸⁰, Jason Coombes⁸⁰, Johnathan M. Evans⁸⁰, Laia Fina⁸⁰, Bree Gatica-Wilcox⁸⁰, Lauren Gilbert⁸⁰, Lee Graham⁸⁰, Jessica Hey⁸⁰, Ember Hilvers⁸⁰, Sophie Jones⁸⁰, Hannah Jones⁸⁰, Sara Kumziene-Summerhayes⁸⁰, Caoimhe McKerr⁸⁰, Jessica Powell⁸⁰, Georgia Pugh⁸⁰, Sarah Taylor⁸⁰, Alexander J. Trotter⁸¹, Charlotte A. Williams¹⁰², Leanne M. Kermack¹⁰⁸, Benjamin H. Foulkes¹¹⁵, Marta Gallis¹¹⁵, Hailey R. Hornsby¹¹⁵, Stavroula F. Louka¹¹⁵, Manoj Pohare¹¹⁵, Paige Wolverson¹¹⁵, Peijun Zhang¹¹⁵, George MacIntyre-Cockett¹²¹, Amy Trebes¹²¹, Robin J. Moll¹²², Lynne Ferguson⁵, Emily J. Goldstein⁵, Alasdair Maclean⁵, Rachael Tomb⁵, Igor Starinski³, Laura Thomson¹²³, Joel Southgate^{58,80}, Moritz U. G. Kraemer²⁹, Jayna Raghwan²⁹, Alex E. Zarebski²⁹, Olivia Boyd⁴⁵, Lily Geidelberg⁴⁵, Chris J. Illingworth¹¹, Chris Jackson¹¹, David Pascall¹¹, Sreenu Vattipally³, Timothy M. Freeman¹¹⁵, Sharon N. Hsu¹¹⁵, Benjamin B. Lindsey¹¹⁵, Keith James¹²², Kevin Lewis¹²², Gerry Tonkin-Hill¹²², Jaime M. Tovar-Corona¹²², MacGregor Cox²⁶, Khalil Abudahab^{20,122}, Mirko Menegazzo²⁰, Ben E. W. Taylor^{20,122}, Corin A. Yeats²⁰, Afrida Mukaddas³, Derek W. Wright³, Leonardo de Oliveira Martins⁸¹, Rachel Colquhoun¹¹⁰, Verity Hill¹¹⁰, Ben Jackson¹¹⁰, J. T. McCrone¹¹⁰, Nathan Medd¹¹⁰, Emily Scher¹¹⁰, Jon-Paul Keatley¹²², Tanya Curran⁵⁹, Sian Morgan¹⁷, Patrick Maxwell²⁶, Ken Smith²⁶, Sahar Eldirdiri²⁷, Anita Kenyon²⁷, Alison H. Holmes^{44,63}, James R. Price^{44,63}, Tim Wyatt⁷⁵, Alison E. Mather⁸¹, Timofey John A. SkvortsovHartley^{83,102}, Martyn Guest⁵⁸, Christine Kitchen⁵⁸, Ian Merrick⁵⁸, Robert Munn⁵⁸, Beatrice Bertolusso³⁹, Jessica Lynch³⁹, Gabrielle Vernet³⁹, Stuart Kirk⁴⁰, Elizabeth Wastnedge⁶², Rachael Stanley⁶⁴, Giles Idle⁷⁰, Declan T. Bradley^{75,83}, Jennifer Poyner⁸⁵, Matilde Mori¹¹⁶, Owen Jones⁵⁸, Victoria Wright²⁴, Ellena Brooks²⁶, Carol M. Churcher²⁶, Mireille Fragakis²⁶, Katerina Galai^{26,76}, Andrew Jermy²⁶, Sarah Judges²⁶, Georgina M. McManus²⁶, Kim S. Smith²⁶, Elaine Westwick²⁶, Stephen W. Attwood²⁹, Frances Bolt^{44,63}, Alisha Davies⁸⁰, Elen De Lacy⁸⁰, Fatima Downing⁸⁰, Sue Edwards⁸⁰, Lizzie Meadows⁸¹, Sarah Jeremiah¹⁰³, Nikki Smith¹¹⁵, Luke Foulser¹²², Themoula Charalampous^{18,52}, Amita Patel¹⁸, Louise Berry²¹, Tim Boswell²¹, Vicki M. Fleming²¹, Hannah C. Howson-Wells²¹, Amelia Joseph²¹, Manjinder Khakh²¹, Michelle M. Lister²¹, Paul W. Bird²², Karlie Fallon²², Thomas Helmer²², Claire L. McMurray²², Mina Odedra²², Jessica Shaw²², Julian W. Tang²², Nicholas J. Willford²², Victoria Blakey²³, Veena Raviprakash²³, Nicola Sheriff²³, Lesley-Anne Williams²³, Theresa Feltwell²⁶, Luke Bedford³², James S. Cargill³³, Warwick Hughes³³, Jonathan Moore³⁴, Susanne Stonehouse³⁴, Laura Atkinson³⁵, Jack C. D. Lee³⁵, Divya Shah³⁵, Adela Alcolea-Medina^{38,118}, Natasha Ohemeng-Kumi^{38,118}, John Ramble^{38,118}, Jasveen Sehmi^{38,118}, Rebecca Williams³⁹, Wendy Chatterton⁴⁰, Monika Pusok⁴⁰, William Everson⁴³, Anibolina Castigador⁵⁰, Emily Macnaughton⁵⁰, Kate El Bouzidi⁵¹, Temi Lampejo⁵¹, Malur Sudhanva⁵¹, Cassie Breen⁵³, Graciela Sluga⁵⁴, Shazaad S. Y. Ahmad^{55,76}, Ryan P. George⁵⁵, Nicholas W. Machin^{55,76}, Debbie Binns⁵⁶, Victoria James⁵⁶, Rachel Blacow⁶¹, Lindsay Coupland⁶⁴, Louise Smith⁶⁵, Edward Barton⁶⁶, Debra Padgett⁶⁶, Garren Scott⁶⁶, Aidan Cross⁶⁷, Mariyam Mirfenderesky⁶⁷, Jane Greenaway⁶⁸, Kevin Cole⁷⁰, Phillip Clarke⁷³, Nichola Duckworth⁷³, Sarah Walsh⁷³, Kelly Bicknell⁷⁴, Robert Impey⁷⁴, Sarah Wyllie⁷⁴, Richard Hopes⁷⁶, Chloe Bishop⁷⁸, Vicki Chalker⁷⁸, Ian Harrison⁷⁸, Laura Gifford⁸⁰, Zoltan Molnar⁸³, Cressida Auckland⁸⁵, Cariad Evans^{91,115}, Kate Johnson^{91,115}, David G. Partridge^{91,115}, Mohammad Raza^{91,115}, Paul Baker⁹², Stephen Bonner⁹², Sarah Essex⁹², Leanne J. Murray⁹², Andrew I. Lawton⁹³, Shirelle Burton-Fanning⁹⁵, Brendan A. I. Payne⁹⁵, Sheila Waugh⁹⁵, Andrea N. Gomes⁹⁷, Maimuna Kimuli⁹⁷, Darren R. Murray⁹⁷, Paula Ashfield⁹⁸, Donald Dobie⁹⁸, Fiona Ashford⁹⁹, Angus Best⁹⁹, Liam Crawford⁹⁹, Nicola Cumley⁹⁹, Megan Mayhew⁹⁹, Oliver Megram⁹⁹, Jeremy Mirza⁹⁹, Emma Moles-Garcia⁹⁹, Benita Percival⁹⁹, Megan Driscoll¹⁰², Leah Ensell¹⁰², Helen L. Lowe¹⁰², Laurentiu Maftel¹⁰², Matteo Mondani¹⁰², Nicola J. Chaloner¹⁰⁵, Benjamin J. Cogger¹⁰⁵, Lisa J. Easton¹⁰⁵, Hannah Huckson¹⁰⁵, Jonathan Lewis¹⁰⁵, Sarah Lowdon¹⁰⁵, Cassandra S. Malone¹⁰⁵, Florence Munemo¹⁰⁵, Manasa Mutingwende¹⁰⁵, Roberto Nicodemi¹⁰⁵, Olga Podplomyk¹⁰⁵, Thomas Somassa¹⁰⁵, Andrew Beggs¹⁰⁶, Alex Richter¹⁰⁶, Claire Cormie¹⁰⁸, Joana Dias¹⁰⁸, Sally Forrest¹⁰⁸, Ellen E. Higginson¹⁰⁸, Mailis Maes¹⁰⁸, Jamie Young¹⁰⁸, Rose K. Davidson¹⁰⁹, Kathryn A. Jackson¹¹³, Lance Turtle¹¹³, Alexander J. Keeley¹¹⁵, Jonathan Ball¹¹⁹, Timothy Byaruhanga¹¹⁹, Joseph G. Chappell¹¹⁹, Jayasree Dey¹¹⁹, Jack D. Hill¹¹⁹, Emily J. Park¹¹⁹, Arezou Fanaie¹²⁰, Rachel A. Hilson¹²⁰, Geraldine Yaze¹²⁰, Stephanie Lo¹²², Safiah Affi¹⁷, Robert Beer¹⁷, Joshua Maksimovic¹⁷, Kathryn McCluggage¹⁷, Karla Spellman¹⁷, Catherine Bresner⁵⁸, William Fuller⁵⁸, Angela Marchbank⁵⁸, Trudy Workman⁵⁸, Ekaterina

Shelest^{19,87}, Johnny Debebe²⁴, Fei Sang²⁴, Marina Escalera Zamudio²⁹, Sarah Francois²⁹, Bernardo Gutierrez²⁹, Tetyana I. Vasylyeva²⁹, Flavia Flaviani³⁷, Manon Ragonnet-Cronin⁴⁵, Katherine L. Smollett², Alice Broos³, Daniel Mair³, Jenna Nichols³, Kyriaki Nomikou³, Lily Tong³, Ioulia Tsatsani³, Sarah O'Brien⁶⁰, Steven Rushton⁶⁰, Roy Sanderson⁶⁰, Jon Perkins⁶¹, Seb Cotton⁶², Abbie Gallagher⁶², Elias Allara^{76,108}, Clare Pearson^{76,108}, David Bibby⁷⁸, Gavin Dabrera⁷⁸, Nicholas Ellaby⁷⁸, Eileen Gallagher⁷⁸, Jonathan Hubb⁷⁸, Angie Lackenby⁷⁸, David Lee⁷⁸, Nikos Manesis⁷⁸, Tamyo Mbisa⁷⁸, Steven Platt⁷⁸, Katherine A. Twohig⁷⁸, Mari Morgan⁸⁰, Alp Aydin⁸¹, David J. Baker⁸¹, Ebenezer Foster-Nyarko⁸¹, Sophie J. Prosolek⁸¹, Steven Rudder⁸¹, Chris Baxter⁸³, SÁlvia F. Carvalho⁸³, Deborah Lavin⁸³, Arun Mariappan⁸³, Clara Radulescu⁸³, Aditi Singh⁸³, Miao Tang⁸³, Helen Morcrette⁸⁵, Nadua Bayzid¹⁰², Marius Cotic¹⁰², Carlos E. Balcazar¹¹⁰, Michael D. Gallagher¹¹⁰, Daniel Maloney¹¹⁰, Thomas D. Stanton¹¹⁰, Kathleen A. Williamson¹¹⁰, Robin Manley¹¹¹, Michelle L. Michelsen¹¹¹, Christine M. Sambles¹¹¹, David J. Studholme¹¹¹, Joanna Warwick-Dugdale¹¹¹, Richard Eccles¹¹³, Matthew Gemmell¹¹³, Richard Gregory¹¹³, Margaret Hughes¹¹³, Charlotte Nelson¹¹³, Lucille Rainbow¹¹³, Edith E. Vamos¹¹³, Hermione J. Webster¹¹³, Mark Whitehead¹¹³, Claudia Wierzbicki¹¹³, Adrienn Angyal¹¹⁵, Luke R. Green¹¹⁵, Max Whiteley¹¹⁵, Emma Betteridge¹²², Iraad F. Bronner¹²², Ben W. Farr¹²², Scott Goodwin¹²², Stefanie V. Lensing¹²², Shane A. McCarthy^{108,122}, Michael A. Quail¹²², Diana Rajan¹²², Nicholas M. Redshaw¹²², Carol Scott¹²², Lesley Shirley¹²², Scott A. J. Thurston¹²², Will Rowe⁴⁹, Amy Gaskin⁸⁰, Thanh Le-Viet⁸¹, James Bonfield¹²², Jennifer Liddle¹²² & Andrew Whitwham¹²²

¹²Barking, Havering and Redbridge University Hospitals NHS Trust, Romford, UK. ¹³Betsi Cadwaladr University Health Board, North Wales, UK. ¹⁴Blackpool Teaching Hospitals NHS Foundation Trust, Blackpool, UK. ¹⁵Bournemouth University, Poole, UK. ¹⁶Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. ¹⁷Cardiff and Vale University Health Board, Cardiff, UK. ¹⁸Centre for Clinical Infection and Diagnostics Research, Department of Infectious Diseases, Guy's and St Thomas' NHS Foundation Trust, London, UK. ¹⁹Centre for Enzyme Innovation, University of Portsmouth, Portsmouth, UK. ²⁰Centre for Genomic Pathogen Surveillance, University of Oxford, Oxford, UK. ²¹Clinical Microbiology Department, Queens Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham, UK. ²²Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester, UK. ²³County Durham and Darlington NHS Foundation Trust, Durham, UK. ²⁴Deep Seq, School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK. ²⁵Department of Infectious Diseases and Microbiology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ²⁶Department of Medicine, University of Cambridge, Cambridge, UK. ²⁷Department of Microbiology, Kettering General Hospital, Kettering, UK. ²⁸Department of Microbiology, South West London Pathology, London, UK. ²⁹Department of Zoology, University of Oxford, Oxford, UK. ³⁰Division of Virology, Department of Pathology, University of Cambridge, Cambridge, UK. ³¹East Kent Hospitals University NHS Foundation Trust, Kent, UK. ³²East Suffolk and North Essex NHS Foundation Trust, Colchester, UK. ³³East Sussex Healthcare NHS Trust, Seaford, UK. ³⁴Gateshead Health NHS Foundation Trust, Gateshead, UK. ³⁵Great Ormond Street Hospital for Children, NHS Foundation Trust, London, UK. ³⁶Great Ormond Street Institute of Child Health (GOS ICH), University College London (UCL), London, UK. ³⁷Guy's and St. Thomas' Biomedical Research Centre, London, UK. ³⁸Guy's and St. Thomas' NHS Foundation Trust, London, UK. ³⁹Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK. ⁴⁰Health Services Laboratories, London, UK. ⁴¹Heartlands Hospital, Birmingham, UK. ⁴²Hub for Biotechnology in the Built Environment, Northumbria University, Newcastle upon Tyne, UK. ⁴³Hull University Teaching Hospitals NHS Trust, Hull, UK. ⁴⁴Imperial College Healthcare NHS Trust, London, UK. ⁴⁵Imperial College London, London, UK. ⁴⁶Infection Care Group, St George's University Hospitals NHS Foundation Trust, London, UK. ⁴⁷Institute for Infection and Immunity, St George's University of London, London, UK. ⁴⁸Barts Health NHS Trust, London, UK. ⁴⁹Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK. ⁵⁰Isle of Wight NHS Trust, Newport, UK. ⁵¹King's College Hospital NHS Foundation Trust, London, UK. ⁵²King's College London, London, UK. ⁵³Liverpool Clinical Laboratories, Liverpool, UK. ⁵⁴Maidstone and Tunbridge Wells NHS Trust, Tunbridge Wells, UK. ⁵⁵Manchester University NHS Foundation Trust, Manchester, UK. ⁵⁶Microbiology Department, Buckinghamshire Healthcare NHS Trust, Amersham, UK. ⁵⁷Microbiology, Royal Oldham Hospital, Oldham, UK. ⁵⁸Cardiff University, Cardiff, UK. ⁵⁹Belfast Health & Social Care Trust, Belfast, UK. ⁶⁰Newcastle University, Newcastle upon Tyne, UK. ⁶¹NHS Greater Glasgow and Clyde, Glasgow, UK. ⁶²NHS Lothian, Edinburgh, UK. ⁶³NIHR Health Protection Research Unit in HCAI and AMR, Imperial College London, London, UK. ⁶⁴Norfolk and Norwich University Hospitals NHS Foundation Trust, Norwich, UK. ⁶⁵Norfolk County Council, Norfolk, UK. ⁶⁶North Cumbria Integrated Care NHS Foundation Trust, Cumbria, UK. ⁶⁷North Middlesex University Hospital NHS Trust, London, UK. ⁶⁸North Tees and Hartlepool NHS Foundation Trust, Durham, UK. ⁶⁹North West London Pathology, London, UK. ⁷⁰Northumbria Healthcare NHS Foundation Trust, Tyneside, UK. ⁷¹Northumbria University, Newcastle upon Tyne, UK. ⁷²NU-OMICS, Northumbria University, Newcastle upon Tyne, UK. ⁷³Path Links, Northern Lincolnshire and Goole NHS Foundation Trust, Lincolnshire, UK. ⁷⁴Portsmouth Hospitals University NHS Trust, Portsmouth, UK. ⁷⁵Public Health Agency, Belfast, Northern Ireland, UK. ⁷⁶Public Health England, London, UK. ⁷⁷Public Health England, Cambridge, UK. ⁷⁸Public Health England, Colindale, UK. ⁷⁹Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ⁸⁰Public Health Wales, Wales, UK. ⁸¹Quadram Institute Bioscience, Norwich, UK. ⁸²Queen Elizabeth Hospital, Birmingham, UK. ⁸³Queen's University Belfast, Belfast, UK. ⁸⁴Royal Brompton and Harefield Hospitals, Uxbridge, UK. ⁸⁵Royal Devon and Exeter NHS Foundation Trust, Exeter, UK. ⁸⁶Royal Free London NHS Foundation Trust, London, UK.

⁸⁷School of Biological Sciences, University of Portsmouth, Portsmouth, UK. ⁸⁸School of Health Sciences, University of Southampton, Southampton, UK. ⁸⁹School of Medicine, University of Southampton, Southampton, UK. ⁹⁰School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth, UK. ⁹¹Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. ⁹²South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK. ⁹³Southwest Pathology Services, Taunton, UK. ⁹⁴Swansea University, Swansea, UK. ⁹⁵The Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. ⁹⁶The Queen Elizabeth Hospital King's Lynn NHS Foundation Trust, King's Lynn, UK. ⁹⁷The Royal Marsden NHS Foundation Trust, London, UK. ⁹⁸The Royal Wolverhampton NHS Trust, Cannock, UK. ⁹⁹Turnkey Laboratory, University of Birmingham, Birmingham, UK. ¹⁰⁰University College London Division of Infection and Immunity, London, UK. ¹⁰¹University College London Hospital Advanced Pathogen Diagnostics Unit, London, UK. ¹⁰²University College London Hospitals NHS Foundation Trust, London, UK. ¹⁰³University Hospital Southampton NHS Foundation Trust, Southampton, UK. ¹⁰⁴University Hospitals Dorset NHS Foundation Trust, Poole, UK. ¹⁰⁵University Hospitals Sussex NHS Foundation Trust, Sussex, UK. ¹⁰⁶University of Birmingham, Birmingham, UK. ¹⁰⁷University of Brighton, Brighton, UK. ¹⁰⁸University of Cambridge, Cambridge, UK. ¹⁰⁹University of East Anglia, Norwich, UK. ¹¹⁰University of Edinburgh, Edinburgh, UK. ¹¹¹University of Exeter, Exeter, UK. ¹¹²University of Kent, Kent, UK. ¹¹³University of Liverpool, Liverpool, UK. ¹¹⁴University of Oxford, Oxford, UK. ¹¹⁵University of Sheffield, Sheffield, UK. ¹¹⁶University of Southampton, Southampton, UK. ¹¹⁷University of St Andrews, St Andrews, UK. ¹¹⁸Viapath, Guy's and St Thomas' NHS Foundation Trust, and King's College Hospital NHS Foundation Trust, London, UK. ¹¹⁹Virology, School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK. ¹²⁰Watford General Hospital, Watford, UK. ¹²¹Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ¹²²Wellcome Sanger Institute, Hinxton, UK. ¹²³Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ¹²⁴Whittington Health NHS Trust, London, UK.