Check for updates

Sequencing analysis of SARS-CoV-2 cases in Slovenian long-term care facilities to support outbreak control

Rok Kogoj[1], Manja Grašek[2], Alen Suljič[1], Samo Zakotnik[1], Doroteja Vlaj[1], Kaja Kotnik Koman[1], Mario Fafangel[2], Miroslav Petrovec[1], Tatjana Avšič-Županc[1] and Misa Korva[1]*

[1]Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia, [2]Communicable Diseases Center, National Institute of Public Health, Ljubljana, Slovenia

Introduction: Residents of long-term care facilities (LTCFs) are at high risk of morbidity and mortality due to COVID-19, especially when new variants of concern (VOC) emerge. To provide intradisciplinary data in order to tailor public health interventions during future epidemics, available epidemiologic and genomic data from Slovenian LTCFs during the initial phases of the COVID-19 pandemic was analyzed.

Methods: The first part of the study included SARS-CoV-2 reverse-transcription Real-Time PCR (rtRT-PCR) positive LTCF residents, from 21 facilities with COVID-19 outbreaks occurring in October 2020. The second part of the study included SARS-CoV-2 rtRT-PCR positive LTCF residents and staff between January and April 2021, when VOC Alpha emerged in Slovenia. Next-generation sequencing (NGS) was used to acquire SARS-CoV-2 genomes, and lineage determination. In-depth phylogenetic and mutational profile analysis were performed and coupled with available field epidemiological data to assess the dynamics of SARS-CoV-2 introduction and transmission.

Results: 370/498 SARS-CoV-2 positive residents as well as 558/699 SARS-CoV-2 positive residents and 301/358 staff were successfully sequenced in the first and second part of the study, respectively. In October 2020, COVID-19 outbreaks in the 21 LTCFs were caused by intra-facility transmission as well as multiple independent SARS-CoV-2 introductions. The Alpha variant was confirmed in the first LTCF resident approximately 1.5 months after the first Alpha case was identified in Slovenia. The data also showed a slower replacement of existing variants by Alpha in residents compared to staff and the general population.

Discussion: Multiple SARS CoV-2 introductions as well as intra-facility spreading impacted disease transmission in Slovenian LTCFs. Timely implementation of control measures aimed at limiting new introductions while controlling in-facility transmission are of paramount importance, especially as new VOCs emerge. Sequencing, in conjunction with epidemiological data, can facilitate the determination of the need for future improvements in control measures to protect LTCF residents from COVID-19 or other respiratory infections.

KEYWORDS

SARS-CoV-2, next-generation sequencing, long-term care facility, outbreak control, phylogeny

1 Introduction

Since the emergence of SARS-CoV-2, various types of long-term care facilities (LTCFs) have experienced the majority of the COVID-19 burden. In the United States alone, more than 2 million documented cases have resulted in more than 160,000 deaths among residents and nearly 3,000 deaths among staff in LTCFs (1, 2). A similar situation has been observed in other parts of the world (3), including South America (4), Canada (5), the European Union (6–9), England (10), China (11), and Australia (12). The combination of vulnerable populations, staff shortages due to sickness leave, insufficient resources for rapid and accurate testing, and lack of personal protective equipment has created a situation referred to as the "perfect storm" (13).

Non-pharmaceutical measures such as wearing masks, physical distancing, quarantine/isolation, frequent hand washing, wiping surfaces with disinfectants, not touching the eyes, nose, and mouth, and sneezing or coughing into elbows have been used to quell this storm and have played an important role in slowing the spread of SARS-CoV-2 in these settings (14). Non-pharmaceutical measures have been implemented in LTCFs around the world in various combinations and strategies depending on the pandemic wave. Most commonly, a combination of wearing personal protective equipment and regular screening of residents, staff, and visitors, regardless of their respiratory symptoms, has been used (15). An immense step forward in the fight against COVID-19 was the vaccines that were developed with unprecedented effort and became available toward the end of 2020 (16). In most EU/EEA countries, COVID-19 vaccination programs have prioritized LTCF residents and identified them as one of the main focus groups for immunization. With the gradual increase in vaccination rates among LTCF residents, there has been a significant reduction in morbidity and mortality in this population (17). However, during the COVID-19 pandemic, several variants of SARS-CoV-2 emerged, some of which were classified as variants of concern (VOC) due to greater transmissibility, ability to cause more severe disease and higher mortality, or enhanced ability to evade natural or vaccine-derived immunity (18). In already high-risk environments such as LTCFs, these variants pose an even greater threat to the wellbeing of the residents while also posing a greater challenge to effective containment.

Next,-generation sequencing (NGS) has proven to be a powerful tool to generate data in the investigation of the COVID-19 pandemic and to contribute to the scientific understanding and public health response (19). Alongside its numerous advantages, it can assist in the investigation of SARS-CoV-2 transmission dynamics in healthcare settings, which can reveal clusters of infection (20, 21). However, NGS alone might not be enough to understand the directionality of spread or low genomic diversity indicating possible spurious transmission events (22).

The aim of this study was to interpret genomic sequencing data obtained in LTCFs to determine virus introduction pathways and to support future mitigation measures, especially in the case of the emergence of viral variants. The first part of the study assessed whether one or multiple introductions of SARS-CoV-2 occurred within the respective LTCFs during October 2020 (the second pandemic wave). Furthermore, we performed a detailed complete genome analysis of SARS-CoV-2 cases in one LTCF at the onset of its first outbreak to increase the resolution of the results. In the second part of the study, we aimed to determine how early VOC Alpha emerged in LTCF residents and to describe the substitution of existing lineages with the more transmissible Alpha variant in these high-risk environments, where more stringent measures were in place compared to the general population.

2 Materials and methods

2.1 Study population and sample selection

In this study, residents and staff from Slovenian LTCFs were included. The vast majority (over 90% of included residents and staff) were from LTCFs that provide care for the older adult, and the minority (under 10% of included residents and staff) were from LTCFs that provide institutional care for adult individuals with physical or mental developmental disabilities, mental health issues, sensory impairments and mobility disorders. The study was divided into two parts.

In the first part, all LTCF residents that tested positive for SARS-CoV-2 in October 2020 and whose samples were analyzed by the Institute of Microbiology and Immunology (IMI) of the Faculty of Medicine, University of Ljubljana, were included. Residents also had to be part of an outbreak in a LTCF, defined as at least two confirmed COVID-19 cases within 14 days. Following these criteria, residents from 21 LTCFs were included. In addition, one specific LTCF was selected in which SARS-CoV-2–positive residents were detected in the period of 1 month without known prior SARS-CoV-2 infections and before vaccination was available.

In the second part of the study, all LTCF residents and staff that tested positive for SARS-CoV-2 from January to April 2021 and whose samples were analyzed by the IMI were included. This included residents from 45 LTCFs and staff from approximately 55 LTCFs (data about the exact work site was not accessible for all of the staff). LTCF staff were defined as anyone that was registered in the national digital administrative database as an employee in a LTCF. A detailed schematic of the study design is shown in Figure 1 and Table 1.

Samples from the selected individuals were nasopharyngeal swabs collected onsite at respective LTCFs by healthcare professionals in Viral Transport Media VTM; (Liofilchem, Roseto degli Abruzzi, Italy) and immediately sent to the IMI via courier for SARS-CoV-2 detection by rtRT-PCR. Samples were tested the same day and then stored at −80°C until the time of the study.

2.2 Sample preparation, library creation, and next-generation sequencing

Total nucleic acids (NA) were isolated from 300 μL of nasopharyngeal swabs collected in VTM. Pure NA eluates were obtained using the OptiPure Viral Auto Plate, Proteinase K Reagent Kit (TANBead, Taoyuan City, Taiwan) on a Maelstrom 9,600 instrument (TANBead, Taoyuan City, Taiwan).

Isolated nucleic acids were prepared for NGS using COVIDSeq Test (Illumina, San Diego, CA, United States) following the
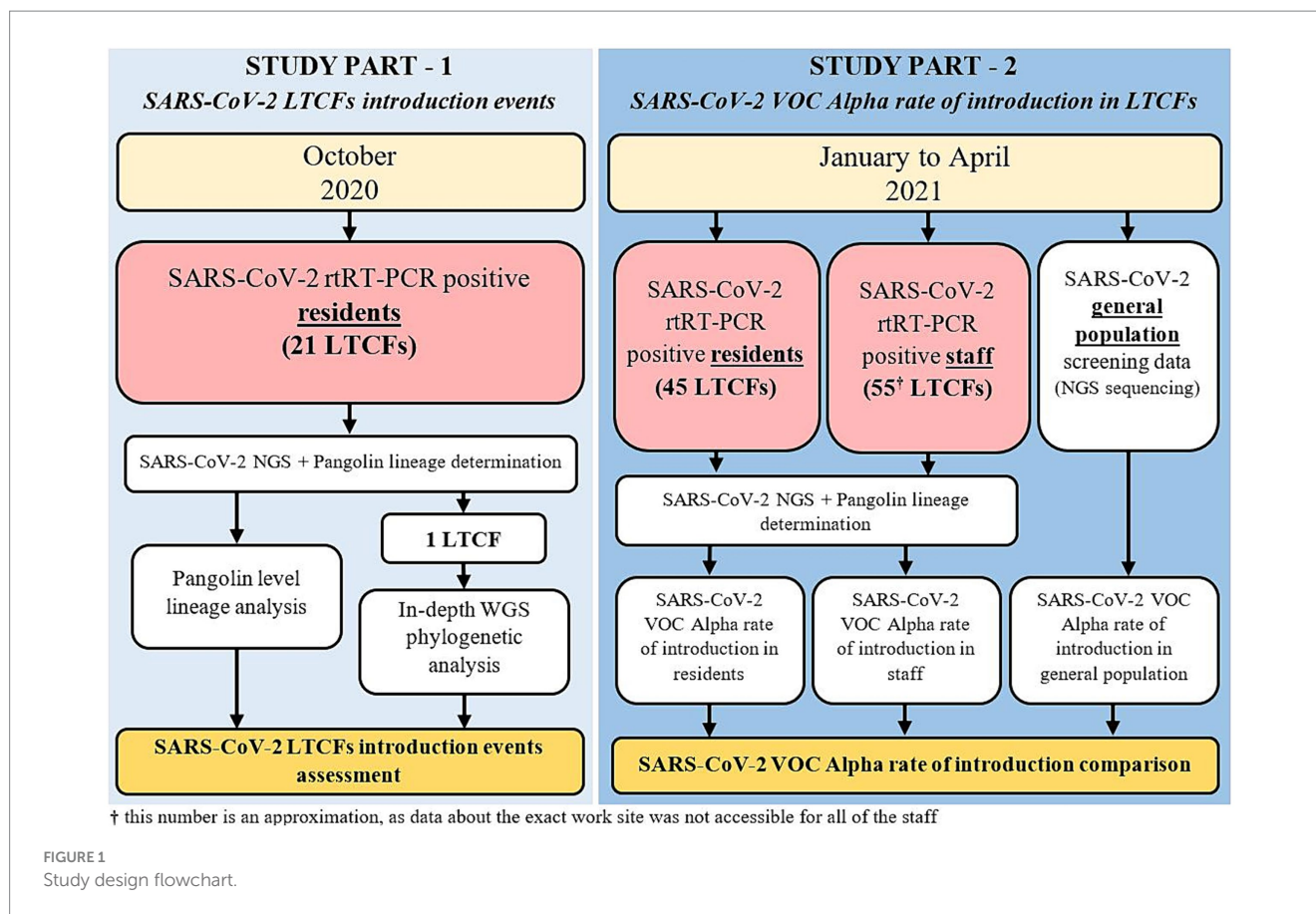

FIGURE 1
Study design flowchart.

TABLE 1 Detailed study characteristics.

Study part	Number of included LTCFs	Time period	Patient category	PCR testing site	SARS-CoV-2 cases
1	21	October 2020	Residents	IMI	498
2	45	January to April 2021	Residents	IMI	699
	55[a]		Staff	IMI	358

[a]This number is an approximation, as data about the exact work site was not accessible for all of the staff.

manufacturer's instructions. The concentration of each library was measured using the Qubit dsDNA HS Assay Kit (Thermo Scientific, Waltham, MA, United States), and fragment size was measured using the Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA, United States). Finally, sequencing was performed using the MiSeq Reagent Kit v3 (600 cycles) on a MiSeq sequencer or the NextSeq 500/550 High Output Kit v2.5 (300 cycles) on a NextSeq 550 sequencer (Illumina, San Diego, CA, United States).

2.3 Data analysis

The in-house developed bioinformatics pipeline consisted of initial trimming of reads with BBDuk (v38.96) and quality control of raw data using FastQC (v0.11.5). The filtered reads were mapped to the reference genome of the Wuhan Hu-1 isolate (NCBI accession number NC_045512.2) using BWA-MEM (v0.7.17-r1188). Subsequent

data processing, including sorting, partner tagging, duplicate tagging, indexing, and coverage depth calculation, was performed using Samtools (v1.9) (23). Variant calls were generated using iVAR (v1.3.1) (24) with the following settings: minimum quality score threshold for base counting set to 20, minimum frequency threshold set to 0.01, and minimum read depth for variant calling set to 10.

Lineages were assigned based on the consensus sequences of near-complete SARS-CoV-2 genomes using the Pangolin tool (v4.2), which implements a dynamic nomenclature (25). The consensus sequences were additionally aligned with MAFFT (v7.490) using a global alignment over 1,000 iterations (26). The construction of the phylogenetic tree based on the multiple sequence alignment was performed using IQ-TREE 2 (v2.2.0) with automatic model selection for tree inference and ultra-fast bootstrap with 1,000 iterations (27). Data analysis was performed with R statistical software version 4.3.0 (R Foundation for Statistical Computing, Vienna, Austria) and Microsoft Excel 2019 version 1801 (Microsoft, Redmond, WA, United States).

TABLE 2 Detailed overall sequencing results for both study periods.

Study period		1	2	
Time period		October 2020	January – April 2021	
No. of assessed LTCFs		21	45	55[a]
Sample provenience		Residents	Residents	Staff
No. PCR confirmed cases		498	699	358
% (No.) sequenced		74.3% (370)	79.8% (558)	84.1% (301)
SARS-CoV-2 variant	A.27	N/A	0.2% (1)	1.0% (3)
	B.1	3.5% (13)	0.2% (1)	0.7% (2)
	B.1.1	0.3% (1)	N/A	0.7% (2)
	B.1.1.7	N/A	5.0% (28)	16.9% (51)
	B.1.1.58	2.9% (11)	11.8% (66)	1.3% (4)
	B.1.1.70	35.6% (133)	3.9% (22)	1.3% (4)
	B.1.146	2.4% (9)	5.4% (30)	4.3% (13)
	B.1.160	32.6% (122)	2.0% (11)	2.3% (7)
	B.1.177	N/A	N/A	0.3% (1)
	B.1.177.28	N/A	N/A	1.0% (3)
	B.1.177.44	N/A	N/A	0.3% (1)
	B.1.258	2.9% (11)	6.1% (34)	4.6% (14)
	B.1.258.17	18.2% (68)	65.2% (364)	63.6% (192)
	B.1.149	N/A	N/A	0.7% (2)
	B.1.531	N/A	N/A	0.3% (1)
	B.1.533	N/A	0.2% (1)	0.3% (1)
	B.1.565	N/A	N/A	0.3% (1)
	C.35	1.6% (6)	N/A	N/A

[a]This number is an approximation, as data about the exact work site was not accessible for all of the staff.

2.4 Ethical approval

This study was performed in accordance with the ethical guidelines for human research, the World Medical Association's Declaration of Helsinki, the Oviedo Convention on Human Rights and Biomedicine, and the Slovenian Code on Medical Deontology. The study was approved by the National Medical Ethics Committee, Ministry of Health, Republic of Slovenia (0120 211/2020/7). The need for informed patient consent was waived, as the manuscript contains no identifying information or images.

3 Results

3.1 Overall distribution of SARS-CoV-2 lineages in LTCFs

Of the 498 COVID-19 positive LTCF residents in the first period of the study, 74.3% (370/498) were successfully sequenced. Of the 699 and 358 COVID-19 positive LTCFs residents and staff in the second period of the study, 79.8% (558/699) and 84.1% (301/358) were successfully sequenced, respectively. Samples for which complete SARS-CoV-2 genome was not obtained, originated from patents with low viral loads and were consequently excluded from further analyses.

In the first study period (October 2020), nine lineages were observed among the residents of the 21 selected LTCFs. The most common lineage was B.1.1.70 with 35.6%, followed by B.1.160 with 32.6% and B.1.258.17 with 18.2%. Six other lineages were present among the residents, which together accounted for the remaining 13.6%.

In the second study period (January to April 2021), 10 lineages were detected among residents and 17 among staff. Lineage B.1.258.17 was the most abundant lineage in both subpopulations, at 65.2 and 63.6% in residents and staff, respectively. Among staff, Alpha is already the second most common lineage at 16.9% in contrast to residents, where Alpha is only the fifth most common lineage (5.0%). Table 2 shows a detailed representation of sequencing results and the lineages detected in both time periods and subpopulations.

3.2 Assessment of SARS-CoV-2 introduction events in LTCFs in October 2020

In October 2020, 21 LTCFs with COVID-19 outbreaks were identified in which there were at least two residents with confirmed SARS-CoV-2 infection in a respective LTCF within 14 days. This included 14/21 (66.6%) LTCFs for which Pangolin SARS-CoV-2
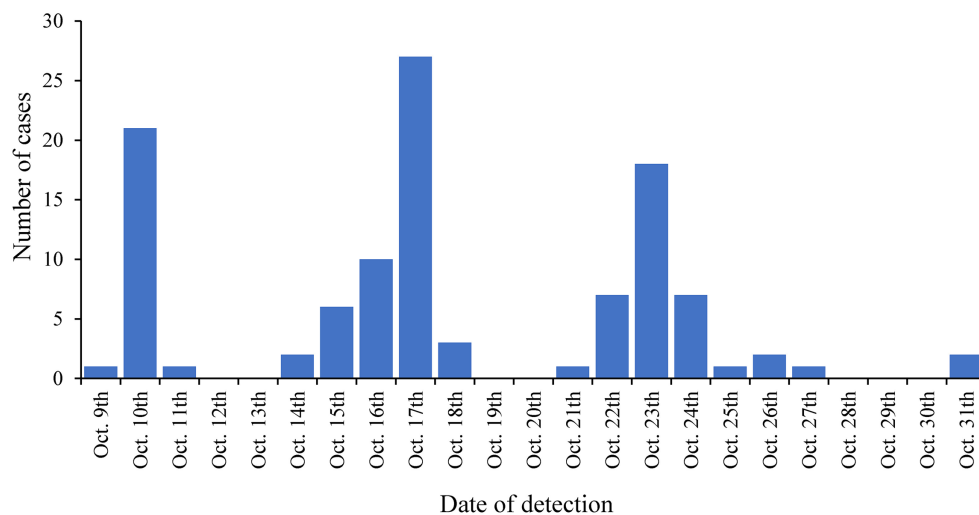
FIGURE 2
Number of sequenced SARS-CoV-2 cases in a specific LTCF per day in October 2020.

lineage designation revealed that the outbreak was caused by one SARS-CoV-2 lineage and 7/21 (33.3%) LTCFs for which multiple variants of SARS-CoV-2 were detected during the outbreak. The co-occurrence of two different lineages were detected in all of these cases: B.1.1.70 and B.1.1, B.1.1.70 and B.1.160, B.1.258.17 and B.1, B.1.258.17 and B.1.258, B.1.1.70 and C.35, B.1.258 and B.1.160, and B.1.258.17 and B.1.1.70.

3.3 Complete genome phylogenetic analysis of the first outbreak in one specific LTCF

The LTCF with the largest outbreak in October 2020 was selected for an in-depth analysis of SARS-CoV-2 complete genome sequences (genomic data are available at GISAID: EPI_SET_231114ud, https://10.55876/gis8.231114ud, accessed on November 14th, 2023). This LTCF was selected because no SARS-CoV-2 cases were detected prior to this period in this facility (an immunologically naive population) and residents were not vaccinated because the vaccine was not yet available. During this period, 145 cases were confirmed among residents, the first on October 9th, 2020, and then throughout the month with peaks on October 17th, 2020, and October 23rd, 2020, as shown in Figure 2. Among these 75.9% (110/145) were successfully sequenced.

Based on the determination of the SARS-CoV-2 lineage by Pangolin, 99.1% (109/110) of the cases collected during the month belonged to lineage B.1.1.70. The only sample from a different lineage (B.1.1) was collected on October 24th, 2020, clearly indicating an independent introduction event already at the Pangolin level (Figure 3). A detailed analysis of the B.1.1.70 cluster, based on complete genome sequences, indicated that there must have been several introductions of this lineage. The first case detected on day 0 already contains more mutations than samples detected at later time points. Moreover, the samples discovered on the same day are scattered across the phylogenetic tree and contain a larger number of mutations than would be possible considering the SARS-CoV-2 mutation fixation rate (28). On the other hand, some pairs are observed that appear to confirm internal transmission due to an appropriate delay in detection (incubation) and do not show any differences in the genome. In addition, some samples collected around the same day suggest a common source of infection because they are genetically identical (Figure 4).

3.4 In-depth mutation profile analysis of the B.1.1.70 subgroup from the selected LTCF

An additional investigation of the mutation profile of the samples of the B.1.1.70 subgroup revealed 24 distinct mutational constellations, suggesting at least five independent introduction events (Figure 5). The distinction between the suspected independent introductions was based on the estimated mutation rate of 1.8×10^{-3} substitutions per base per year, which corresponds to one additional mutation per week (28). There are four sample subgroups (designated by the numbers 1–4) with more than one additional substitution already present in the 1st week of investigation (time delay <7 days), indicating independent introduction events. Another possible introduction event is evident on the 9th day with four additional substitutions (mutation profile 5). All other mutation profiles (designated by number and letter; 1a–1h, 3a–3e) suggest either intra-LTCF transmission or independent introduction events based on the number of new mutations and time delays, but without additional epidemiologic data the source cannot be resolved. The thresholds for suspected individual introduction events are shown in Figure 6.

3.5 VOC Alpha rate of introduction into LTCFs

In the second part of the study (January 2021 – April 2021), a total of 79.8% (558/699) and 84.1% (301/358) of the samples from
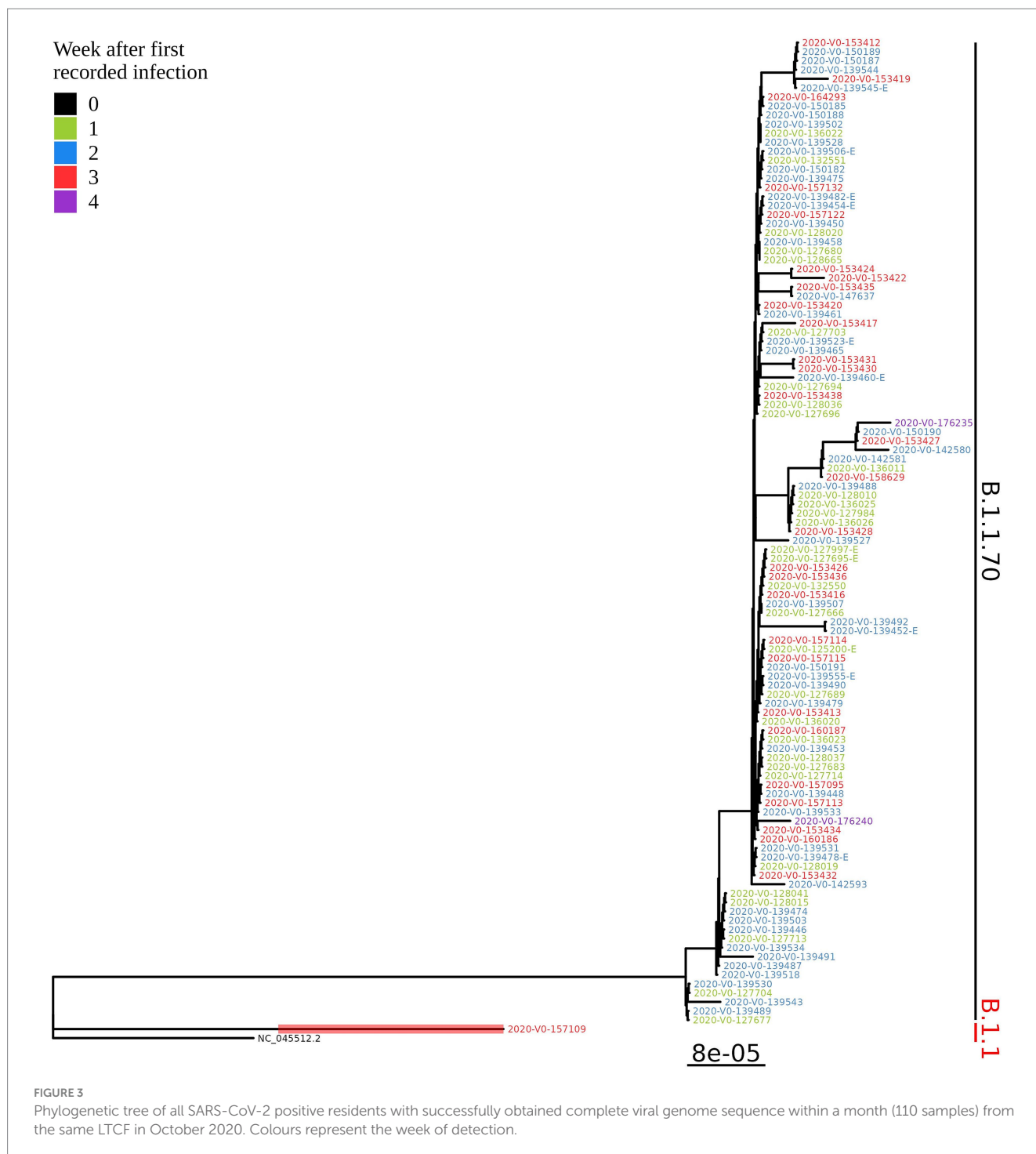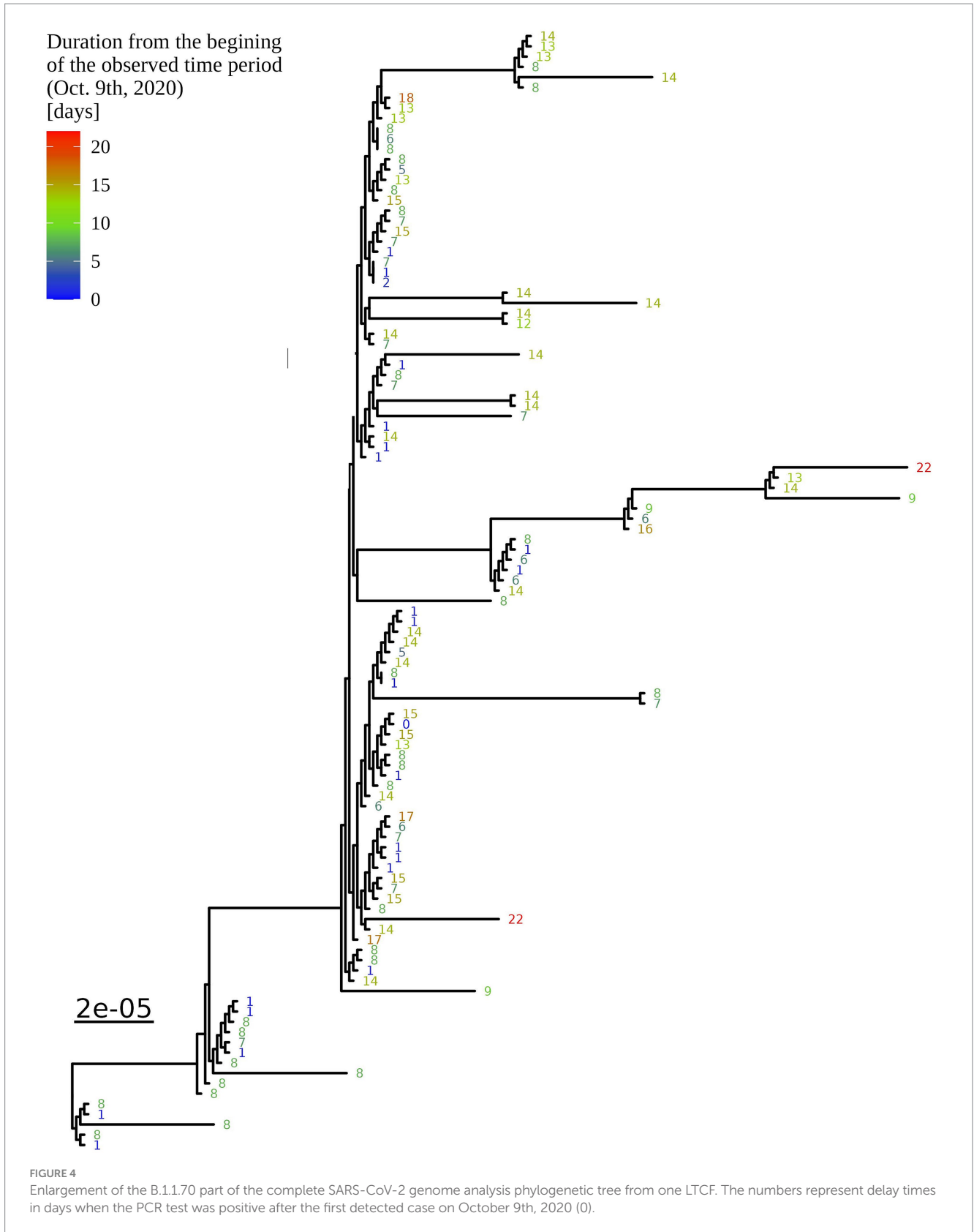
FIGURE 3
Phylogenetic tree of all SARS-CoV-2 positive residents with successfully obtained complete viral genome sequence within a month (110 samples) from the same LTCF in October 2020. Colours represent the week of detection.

residents and staff from selected LTCFs were successfully sequenced, respectively. At the beginning of the selected period, Alpha was not present in the selected LTCF population (residents and staff) or in the general Slovenian population. The first case of the Alpha variant was detected in the general population on January 7th, 2021. The first Alpha cases in LTCF residents enrolled in the study were detected in week 8/2021. After a 2 week period without detection, Alpha was detected again and accounted for 50.0% of the cases detected. The proportion did not increase until the end of the study period, but it actually decreased slightly, even though the number of

positive SARS-CoV-2 tests increased. In contrast, the first Alpha case was detected among LTCFs staff in week 7/2021. After isolated cases in the next 2 weeks and no detection in week 10, Alpha reappeared in week 11. In the following weeks, the prevalence continued to increase until it almost completely replaced the previous lineages. The temporal distribution of detected lineages in LTCF residents and staff from January 2021 to April 2021 is shown in Figure 7.

An additional analysis of the proportion of Alpha cases revealed that the growth rate of Alpha cases at the time of introduction to the

FIGURE 4
Enlargement of the B.1.1.70 part of the complete SARS-CoV-2 genome analysis phylogenetic tree from one LTCF. The numbers represent delay times in days when the PCR test was positive after the first detected case on October 9th, 2020 (0).

LTCFs was similar for all three groups observed (i.e., LTCF residents, LTCF staff, and the general population) until the second half of the period analyzed. Thereafter, the growth rate of Alpha was significantly slower for residents than for the general population and LTCF staff. In

addition, the analysis of the diversity of lineages (number of different lineages present per day) in the three populations showed that, in general population, the diversity of circulating lineages increases at the beginning of the study period and starts decreasing toward the middle
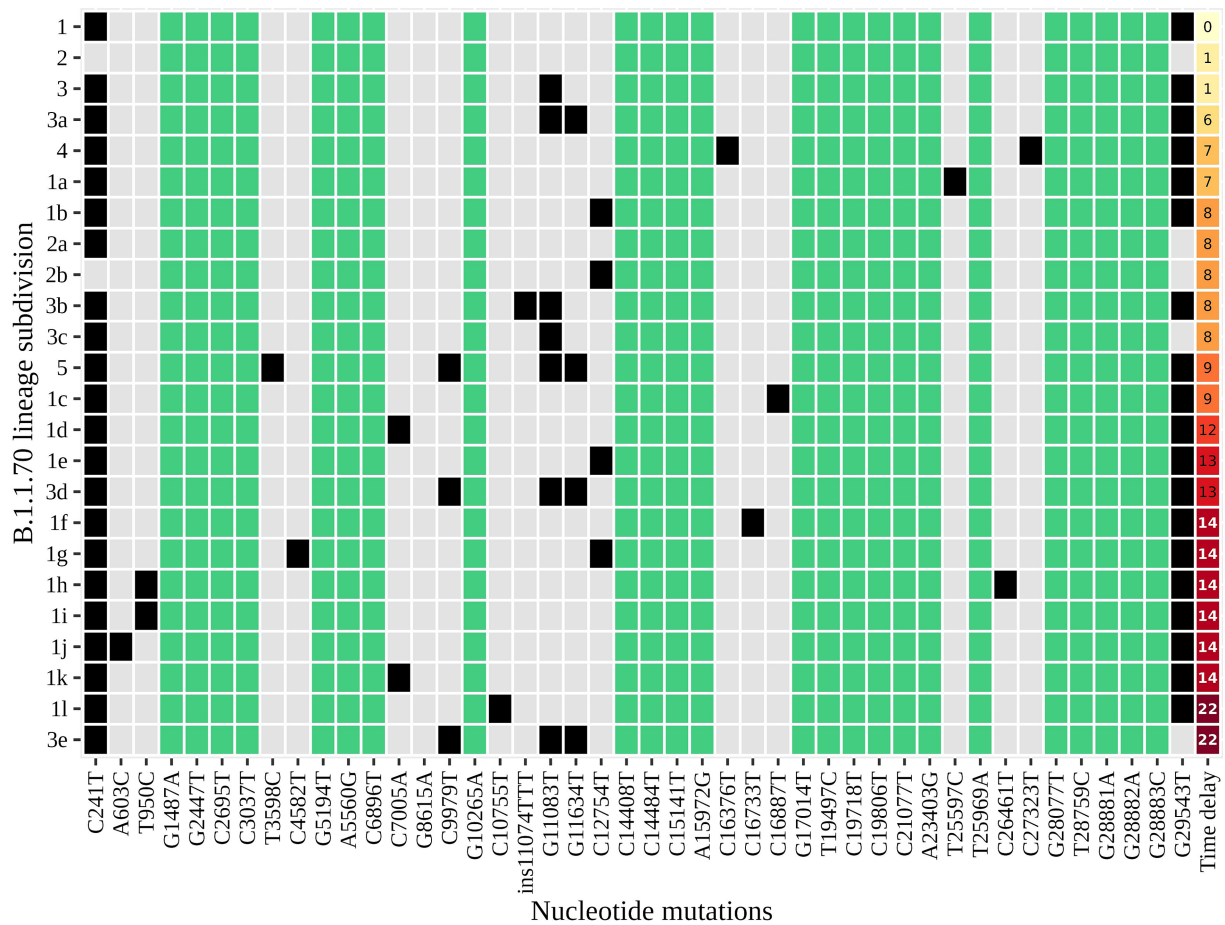
FIGURE 5
The distinct mutation profile of the B.1.1.70 subgroup in one specific LTCF. The y-axis represents the classification beyond Pangolin, with a sequential number denoting a potential independent introduction event and a letter if the mutation profile represents an expected change within a certain timeframe (possible intra-LTCF transmission). Green represents substitutions that are fixed, grey represents substitutions that differ between samples, and black represents the presence of substitutions in a mutation profile. The last column represents the time delay after the first detection of SARS-CoV-2.

of the selected period due to Alpha spread. In contrast, lineage diversity decreases in both LTCF residents and staff right at the beginning of the selected period (Figure 8). The number of SARS-CoV-2 infections in LTCFs residents and staff and the number of SARS-CoV-2 infections in the general Slovenian population during the period observed are presented in Figure 8. In some of the weeks observed, the absolute number of cases among LTCF staff and especially among residents was very low.

4 Discussion

In this retrospective study, the introduction and subsequent spread of SARS-CoV-2 cases in Slovenian LTCFs were analyzed in detail using NGS, with the aim of providing interdisciplinary data in order to tailor public health interventions during future epidemics. Two time periods based on the epidemiological situation in the country were selected: the initial pandemic phase, when several early lineages were co-circulating in the population, and the period of VOC Alpha appearance in Slovenia, followed by a significant surge in cases.

In the first part of the study, COVID-19 outbreaks in selected Slovenian LTCFs in October 2020 were additionally investigated using complete genome sequencing to identify different variants or introductions of SARS-CoV-2 in each specific outbreak. At that time, Slovenian LTCFs had already been instructed to implement a range of containment strategies such as, the use of personal protective equipment, the establishment of separate zones for residents with confirmed or suspected infections and, monitoring the health status of employees. Because the analyses were conducted at the level of individual LTCFs and exact epidemiological data on all employees in these facilities were not readily available and verifiable (e.g., their contact with residents, their employment in one or more LTCFs, their absence from work, etc.), staff were not included in this part of the study. The findings revealed that the outbreaks in some LTCFs were attributable to two SARS-CoV-2 variants, directly indicating at least two independent sources of infections. However, in the majority of LTCFs, lineage determination at the Pangolin level revealed a single variant, which appears to indicate a single introduction event. However, a more in-depth analysis revealed that instances in which a single Pangolin variant was identified as a cause of an outbreak do not necessarily point to a single SARS-CoV-2 introduction event. In fact,
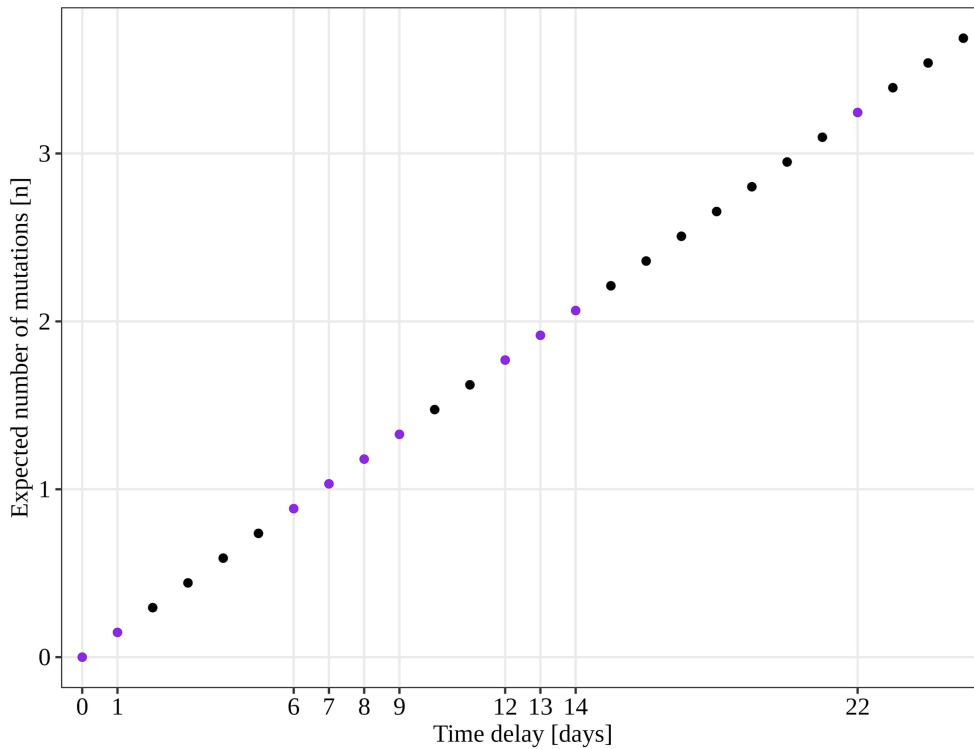
FIGURE 6
Estimated number of SARS-CoV-2 mutations accumulated in a closed system based on 1.8×10^{-3} mutations acquisition per base per year (calculated for the SARS-CoV-2 reference genome NC_045512.2). The purple dots represent the time points observed in our study and the expected number of mutations if only intra LTCF transmission were to occur.
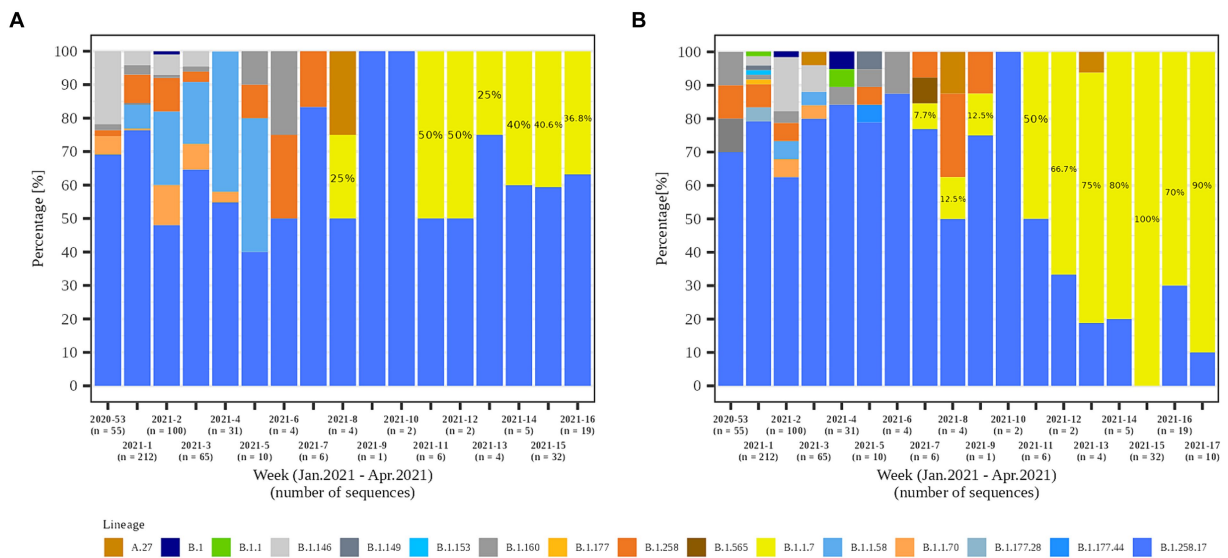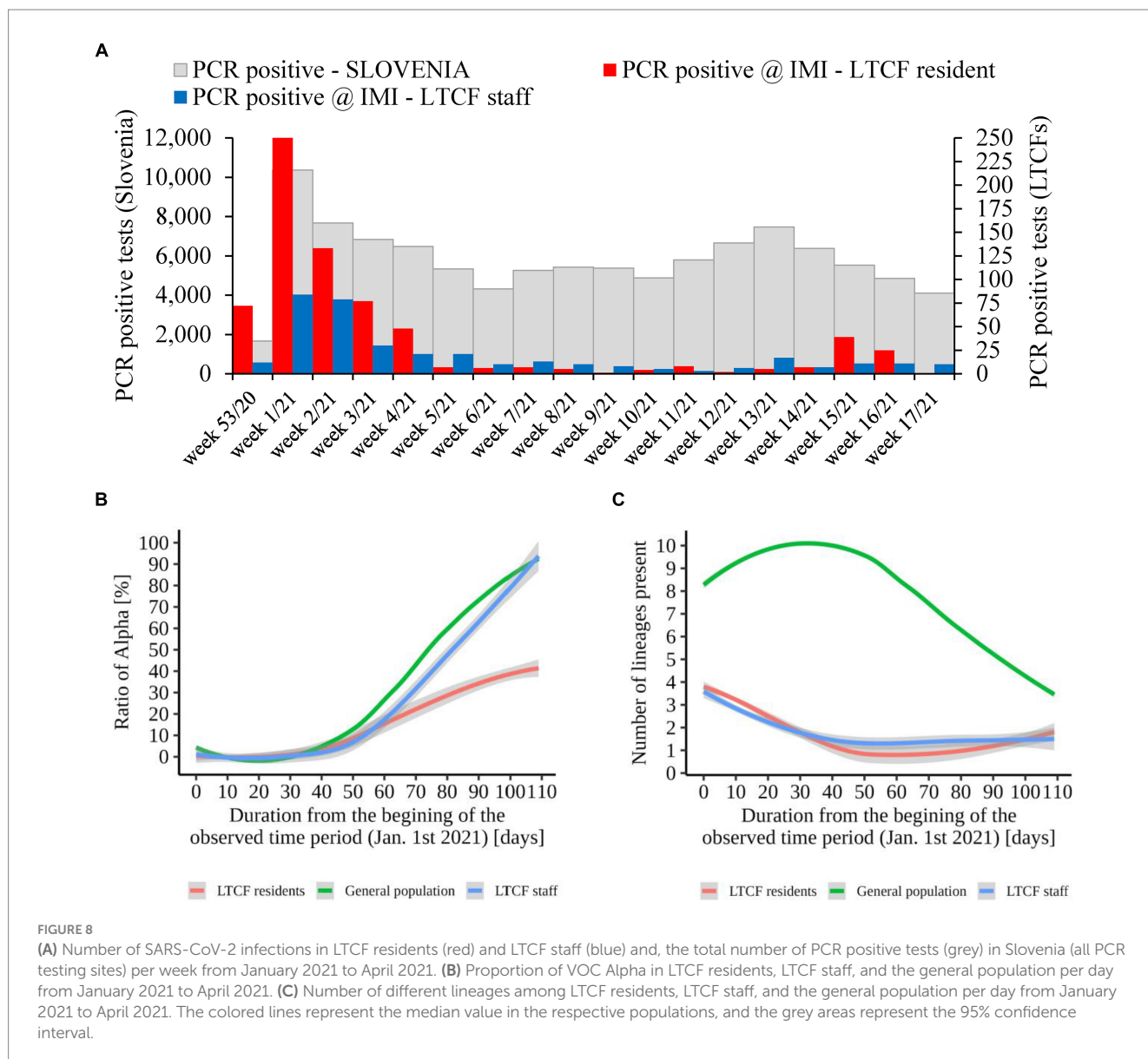

FIGURE 7
Temporal distribution of detected SARS-CoV-2 lineages in **(A)** LTCF residents and **(B)** LTCF staff per week of detection.

complete genome phylogenetic and mutation profiles analysis of samples from the largest LTCF outbreak in October 2020 showed multiple introductions of the SARS-CoV-2 into the LTCF. Such findings would not have been possible based on the Pangolin lineage designation alone. Moreover, comprehensive contact tracing, which

in certain cases could enable the acquisition of such data, is often difficult to carry out in critical situations due to overburdened staff and it tends to be less reliable compared to data acquired faster through sequencing. Obtained data is consistent with some previously published studies that identified several separate introductions of

FIGURE 8
(A) Number of SARS-CoV-2 infections in LTCF residents (red) and LTCF staff (blue) and, the total number of PCR positive tests (grey) in Slovenia (all PCR testing sites) per week from January 2021 to April 2021. **(B)** Proportion of VOC Alpha in LTCF residents, LTCF staff, and the general population per day from January 2021 to April 2021. **(C)** Number of different lineages among LTCF residents, LTCF staff, and the general population per day from January 2021 to April 2021. The colored lines represent the median value in the respective populations, and the grey areas represent the 95% confidence interval.

SARS-CoV-2 within one LTCF (29–32). This highlights the importance of not only implementing strategies to limit the spread of infection within LTCFs, but also taking measures to reduce the likelihood of further introduction of the virus from external sources. Although the potential sources of virus introduction into LTCFs are fairly well known (32), based on our results, measures to prevent such instances could include not only the implementation of procedures (33), but also rapid molecular diagnostics coupled with timely available sequencing, as it enables acquiring additional information for future tailoring of public health measures in such environments. Moreover, in the future, sequencing with rapid turnover time could allow for individualized interventions for specific VOCs, if such differentiated actions would have evidence-based background.

The second part of the study aimed to study the substitution of the predominant variants by the emerging, more transmissible VOC Alpha and to ascertain the earliest instance of its introduction in the LTCFs. We also wanted to determine whether a new VOC could spread faster due to the closed environment of LTCFs and the more

immunologically vulnerable population. Therefore, the period from January to April 2021 was chosen to cover the time from the discovery of the Alpha variant in Slovenia to the time it became dominant in the general population. This timeframe allowed us to capture the initial cases of the Alpha variant in LTCFs and investigate its subsequent spread. The analysis was not conducted at the level of individual LTCFs, but only by selected population groups as a whole (general population, LTCF residents, and LTCF staff). As expected, the staff cohort exhibited a greater array of different variants than the LTCF residents. This result is not surprising because the former typically have more ongoing interactions outside the facility, exposing them to a wider range of infection opportunities. However, this fact makes employees an important source for the introduction of different variants, which is why they should strictly adhere to all recommended infection prevention and control measures. This is particularly important when a new variant emerges that is more immunologically evasive and can easily re-infect individuals (34). The first case of the Alpha variant in the Slovenian general population was detected

approximately a month and a half prior to its detection in the first LTCF resident. At that time, the proportion of the Alpha variant in the general population was close to 15% (35), which may indicate that existing interventions, non-pharmaceutical measures, and vaccination, as well as immunity from prior infections, have slightly delayed the introduction and spread of this VOC in LTCFs.

The assessment of the VOC Alpha proportion showed an initial similarity between the three cohorts observed during the early stages of the spread of this variant, followed by a significant acceleration of the growth trajectory among the general population and LTCF staff in contrast to LTCF residents. At the end of the period observed, when Alpha was already dominant in the general population and LTCF staff, it still barely reached 40% among residents. Slovenia had a very good outcome in the first wave of the pandemic, with several SARS-CoV-2 B lineages circulating. However, in the late summer of 2020, variant B.1.258.17 emerged and spread rapidly throughout the country, causing high mortality (36), partially also because of higher pathogenicity and fitness of the variant. In addition, the vaccination campaign started at the end of December 2020, when LTCF residents and healthcare professionals were the priority groups for vaccination. Exact data on the immunization coverage of Slovenian LTCF residents in January to April 2021 are not available at a national level because uniform reporting to the National Institute of Public Health (NIPH) started in September 2021. However, a questionnaire conducted by the Association of Social Institutions of Slovenia among LTCFs showed that, by mid-June 2021, the average vaccination coverage among Slovenian LTCF residents with at least one dose was approximately 83%. Based on the later assessment of vaccination coverage by the NIPH, in which the vaccination coverage was broadly comparable between LTCFs, we assume that this was also the case with these questionnaire-based results. In contrast, by mid-June 2021 the proportion of the general population over 18 years old vaccinated with at least one dose and the proportion of fully vaccinated individuals was 46 and 34%, respectively (37). The collective contribution of the high prevalence of prior infections with B.1.258.17 and vaccination efforts among residents together with several implemented mitigation strategies (cohorting residents with confirmed or suspected infections, use of personal protective equipment, control measures that allow safe visits, monitoring of health status of staff and residents, hygiene precautions, cleaning and disinfection, etc.) hindered the introduction and spread of the Alpha variant in LTCFs. The decrease in the number of lineages present in both LTCF residents and staff in the early part of the observation period and the simultaneous increase in the general population may also indicate that this closed environment was at least reasonably well protected from additional introduction of different variants, probably by the same factors as mentioned above.

The findings in both the first and second parts of the study emphasize the importance of effective pharmaceutical and non-pharmaceutical preventive measures to curb the introduction of SARS CoV-2 into LTCFs while controlling the spread of infection within them. This is particularly important because residents of LTCFs are at high risk of severe disease progression and increased mortality due to their underlying health conditions, frailty, or advanced age, in addition to the confined living spaces inherent to these facilities (38, 39). Nonetheless, the implementation of non-pharmaceutical interventions in LTCFs requires caution and careful consideration because certain interventions, although intended to protect life, can significantly affect the overall quality of life of residents (40, 41). For

example, if implemented non-pharmaceutical interventions restrict physical, face-to-face interactions or affect the availability of healthcare services, this may exacerbate the social vulnerability of LTCF residents (33). Thus, should a novel SARS-CoV-2 variant or any other potentially threatening pathogen emerge for which the impact and efficacy of vaccination remain uncertain, LTCF residents should be prioritized for both pharmaceutical and non-pharmaceutical interventions, among which interventions that exert the least negative implications on their quality of life should be prioritized.

The limitations of the study should be considered when interpreting these results. The first part of the study included all LTCF residents whose infection was rtRT-PCR confirmed in October 2020 and who were part of outbreaks, as defined in the Materials and Methods section. However, because certain outbreaks persisted over extended periods of time and spanned months in which a dominant variant was already present, individuals whose infections were confirmed after October 2020 were not included. Therefore, the actual number of different variants and introductions that contributed to a LTCF outbreak could be higher if the entire pool of additional residents and a longer study period had been included. Nevertheless, we believe that the approach employed fulfilled its intended purpose to identify the potential for multiple introductions within a single outbreak. An important limitation of the second part of this study is that the number of confirmed cases among LTCF staff and especially residents included in the study was very low in some of the weeks observed, thus broadening the confidence interval of observed data. However, these numbers still provide an accurate description of the COVID-19 epidemiological situation in LTCFs, as they include all rtRT-PCR confirmed residents and staff of LTCFs, with no additional exclusion criteria. Although the IMI has performed on average nearly half of all SARS CoV-2 PCR tests in the country at any given time during the pandemic it mainly covers the central, western, and southern geographical and statistical regions of Slovenia. Because the LTCF staff included in the second part of the study were defined as all persons that had a positive SARS CoV-2 test in the selected period and were registered in the national digital administrative database as employees in LTCF, these criteria also include all those that were not necessarily involved in the outbreaks or had any contact with residents because these data were not readily available. However, the decision to include staff in this segment, despite the aforementioned limitations, was justified by the fact that, due to the intense workload during this period, most of these LTCFs staff infections would have had some form of impact on the workplace.

Despite the aforementioned limitations, we believe that the data obtained, combined with the additional body of knowledge from the LTCFs, could prove useful in future efforts to control communicable diseases. We have shown that sequencing is a powerful supplementary tool to distinguish between purely internal spread and the occurrence of additional external introductions in the context of outbreaks in closed and high-density environments. In this context, our research supports the use of complete genome sequencing as a valuable tool because these types of data combined with epidemiological data can provide useful information for future public health implications not only limited to COVID-19, but also for other potential threats that may emerge in the future.

As revealed by this study, COVID-19 outbreaks in Slovenian LTCFs have been caused by intra-facility transmission as well as several different SARS-CoV-2 introductions. Substitution of the current variants with VOC Alpha appeared to be slower in LTCF residents than

in LTCF staff or the general population, possibly reflecting the effectiveness of the non-pharmaceutical interventions implemented in combination with the uptake of the vaccine among residents, as well as at least partial immunity due to prior infections and vaccination efforts. Given the substantial impact of COVID-19 on LTCFs, it is imperative to prioritize measures in such facilities and adopt timely and proportionate mitigation strategies to effectively protect the residents from SARS-CoV-2, including new VOC variants and any other potential threats that may emerge in the future. We show that the data obtained by genome sequencing, in conjunction with epidemiological data, can help identify the need for future improvement of additional control measures aimed at protecting LTCF residents from other impeding threats, extending beyond COVID-19.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://gisaid.org/, EPI_SET_231114ud.

Ethics statement

The studies involving humans were approved by National Medical Ethics Committee, Ministry of Health, Republic of Slovenia (0120 211/2020/7). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because the manuscript contains no identifying information or images and the results have no direct impact on patient's management or health.

Author contributions

RK: Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing. MG: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. AS: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing, Software. SZ: Formal analysis, Methodology, Writing – review & editing, Software. DV: Methodology, Writing – review & editing. KKK: Methodology, Writing – review & editing. MF: Conceptualization, Methodology, Resources, Supervision, Writing

– original draft, Writing – review & editing, Project administration. MP: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing, Project administration. TA-Ž: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing, Project administration. MK: Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Validation, Writing – review & editing, Project administration.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Meehan A, Gadbois EA, Brazier JF, Madrigal C, Grabowski DC, Shield RR. "It's not a sprint, it's a marathon": stages of managing a nursing home throughout the pandemic. *J Am Med Dir Assoc*. (2023) 24:1579–1585.e2. doi: 10.1016/j.jamda.2023.04.024

2. Konetzka RT, White EM, Pralea A, Grabowski DC, Mor V. A systematic review of long-term care facility characteristics associated with COVID-19 outcomes. *J Am Geriatr Soc*. (2021) 69:2766–77. doi: 10.1111/jgs.17434

3. Dykgraaf SH, Matenge S, Desborough J, Sturgiss E, Dut G, Roberts L, et al. Protecting nursing homes and long-term care facilities from COVID-19: a rapid review of international evidence. *J Am Med Dir Assoc*. (2021) 22:1969–88. doi: 10.1016/j.jamda.2021.07.027

4. Santana RF, Silva MBD, Marcos DADSR, Rosa CS, Wetzel Junior W, Delvalle R. Nursing recommendations for facing dissemination of COVID-19 in Brazilian nursing homes. *Rev Bras Enferm*. (2020) 73:e20200260. doi: 10.1590/0034-7167-2020-0260

5. Webster P. COVID-19 highlights Canada's care home crisis. *Lancet*. (2021) 397:183. doi: 10.1016/S0140-6736(21)00083-0

6. Araujo O, Sequeira C, Ferre-Grau C, Sousa L. [nursing homes in Portugal during the COVID-19 outbreak: challenges for the future.] Spanish. *Enferm Clin*. (2022) 32:83–91. doi: 10.1016/j.enfcli.2021.09.011

7. Rabilloud M, Riche B, Etard JF, Elsensohn MH, Voirin N, Benet T, et al. COVID-19 outbreaks in nursing homes: a strong link with the coronavirus spread in the

surrounding population, France, march to July 2020. *PLoS One*. (2022) 17:e0261756. doi: 10.1371/journal.pone.0261756

8. Malara A, Noale M, Abbatecola AM, Borselli G, Cafariello C, Fumagalli S, et al. Clinical features of SARS-CoV-2 infection in Italian long-term care facilities: Gero Covid LTCFs observational study. *J Am Med Dir Assoc*. (2022) 23:15–8. doi: 10.1016/j.jamda.2021.10.009

9. Roth A, Feller S, Ruhnau A, Plamp L, Viereck U, Weber K, et al. Characterization of COVID-19 outbreaks in three nursing homes during the first wave in Berlin, Germany. *Sci Rep*. (2021) 11:24441. doi: 10.1038/s41598-021-04115-9

10. Schultze A, Nightingale E, Evans D, Hulme W, Rosello A, Bates C, et al. Mortality among care home residents in England during the first and second waves of the COVID-19 pandemic: an observational study of 4.3 million adults over the age of 65. *Lancet Reg Health Eur*. (2022) 14:100295. doi: 10.1016/j.lanepe.2021.100295

11. Guo Y, Liu X, Deng M, Liu P, Li F, Xie N, et al. Epidemiology of COVID-19 in older persons, Wuhan, China. *Age Ageing*. (2020) 49:706–12. doi: 10.1093/ageing/afaa145

12. Gilbert G. COVID-19 in a Sydney nursing home: a case study and lessons learnt. *Med J Aust*. (2020) 213:393–6e1. doi: 10.5694/mja2.50817

13. Ouslander JG, Grabowski DC. COVID-19 in nursing homes: calming the perfect storm. *J Am Geriatr Soc*. (2020) 68:2153–62. doi: 10.1111/jgs.16784. PMID: 32735036

14. Leung NHL. Transmissibility and transmission of respiratory viruses. *Nat Rev Microbiol*. (2021) 19:528–45. doi: 10.1038/s41579-021-00535-6

15. Wassermann M, Ouslander JG, Lam A, Morley JE, von Preyss-Friedman S, Marco N, et al. Diagnostic testing for SARS-CoV-2 in the nursing facility: recommendations of a Delphi panel of long-term care clinicians. *J Nutr Health Aging*. (2020) 24:538–43. doi: 10.1007/s12603-020-1401-9

16. European Centre for Disease Prevention and Control. *COVID-19 vaccination and prioritisation strategies in the EU/EEA. 2020*. Stockholm: ECDC (2020).

17. European Centre for Disease Prevention and Control. *COVID-19 outbreaks in long-term care facilities in the EU/EEA in the context of current vaccination coverage, 2021*. Stockholm: ECDC (2021).

18. Salehi-Vaziri M, Fazlalipour M, Seyed Khorrami SM, Azadmanesh K, Pouriayevali MH, Jalali T, et al. The ins and outs of SARS-CoV-2 variants of concern (VOCs). *Arch Virol*. (2022) 167:327–44. doi: 10.1007/s00705-022-05365-2

19. Saravanan KA, Panigrahi M, Kumar H, Rajawat D, Nayak SS, Bhushan B, et al. Role of genomics in combating COVID-19 pandemic. *Gene*. (2022) 823:146387. doi: 10.1016/j.gene.2022.146387

20. Aggarwal D, Warne B, Jahun AS, Hamilton WL, Fieldman T, Du Plessis L, et al. Genomic epidemiology of SARS-CoV-2 in a UK university identifies dynamics of transmission. *Nat Commun*. (2021) 13:751. doi: 10.1038/s41467-021-27942-w

21. Paltansing S, Sikkema RS, de Man SJ, Koopmans MPG, Oude Munnink BB, de Man P. Transmission of SARS-CoV-2 among healthcare workers and patients in a teaching hospital in the Netherlands confirmed with whole-genome sequencing. *J Hosp Infect*. (2021) 110:178–83. doi: 10.1016/j.jhin.2021.02.005

22. Villabona-Arenas CJ, Hanage WP, Tully DC. Phylogenetic interpretation during outbreaks requires caution. *Nat Microbiol*. (2020) 5:876–7. doi: 10.1038/s41564-020-0738-5

23. Li H, Durbin R. Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*. (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324

24. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using primal Seq and IVar. *Genome Biol*. (2019) 20:8. doi: 10.1186/s13059-018-1618-7

25. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. (2020) 5:1403–7. doi: 10.1038/s41564-020-0770-5

26. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. (2013) 30:772–80. doi: 10.1093/molbev/mst010

27. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. (2020) 37:1530–4. doi: 10.1093/molbev/msaa015

28. Abbasian MH, Mahmanzar M, Rahimian K, Mahdavi B, Tokhanbigli S, Moradi B, et al. Global landscape of SARS-CoV-2 mutations and conserved regions. *J Transl Med*. (2023) 21:152. doi: 10.1186/s12967-023-03996-w

29. Deng X, Gu W, Federman S, du Plessis L, Pybus OG, Faria NR, et al. Genomic surveillance reveals multiple introductions of SARS-CoV-2 into northern California. *Science*. (2020) 369:582–7. doi: 10.1126/science.abb9263

30. Graham NSN, Junghans C, Downes R, Sendall C, Lai H, McKirdy A, et al. SARS-CoV-2 infection, clinical features and outcome of COVID-19 in United Kingdom nursing homes. *J Infect*. (2020) 81:411–9. doi: 10.1016/j.jinf.2020.05.073

31. Hamilton WL, Tonkin-Hill G, Smith ER, Aggarwal D, Houldcroft CJ, Warne B, et al. COVID-19 genomics consortium UK. Genomic epidemiology of COVID-19 in care homes in the east of England. *eLife*. (2021) 10:e64618. doi: 10.7554/eLife.64618

32. Ladhani SN, Chow JY, Janarthanan R, Fok J, Crawley-Boevey E, Vusirikala A, et al. Investigation of SARS-CoV-2 outbreaks in six care homes in London, April 2020. *EClinicalMedicine*. (2020) 26:100533. doi: 10.1016/j.eclinm.2020.100533

33. European Centre for Disease Prevention and Control. *Increase in fatal cases of COVID-19 among long-term care facility residents in the EU/EEA and the UK*. Stockholm: ECDC (2020).

34. Suljič A, Sočan M, Mrzel M, Lunar MM, Korva M, Štorman A, et al. Milder outcomes of SARS-CoV-2 genetically confirmed reinfections compared to primary infections with the delta variant: a retrospective case-control study. *Front Med (Lausanne)*. (2022) 9:962653. doi: 10.3389/fmed.2022.962653

35. Institute of microbiology and Immunology (IMI). *Poročila sistematičnega genomskega opredeljevanja in spremljanje epidemije SARS-CoV-2 [Reports of the systematic genomic SARS-CoV-2 pandemic surveillance]*. Ljubljana: IMI.

36. Ružić Gorenjec N, Kejžar N, Manevski D, Pohar Perme M, Vratanar B, Blagus R. COVID-19 in Slovenia, from a success story to disaster: what lessons can be learned? *Life*. (2021) 11:1045. doi: 10.3390/life11101045

37. National institute of Public Health (NIJZ). *Cepljenje proti COVID-19 v Sloveniji. [Vaccination against COVID-19 in Slovenia]*. Ljubljana: NIJZ.

38. Levin AT, Jylhävä J, Religa D, Shallcross L. COVID-19 prevalence and mortality in longer-term care facilities. *Eur J Epidemiol*. (2022) 37:227–34. doi: 10.1007/s10654-022-00861-w

39. Resciniti NV, Fuller M, Sellner J, Lohman MC. COVID-19 incidence and mortality among long-term care facility residents and staff in South Carolina. *J Am Med Dir Assoc*. (2021) 22:2026–31.e1. doi: 10.1016/j.jamda.2021.08.006

40. Van der Roest HG, Prins M, van der Velden C, Steinmetz S, Stolte E, van Tilburg TG, et al. The impact of COVID-19 measures on well-being of older long-term care facility residents in the Netherlands. *J Am Med Dir Assoc*. (2020) 21:1569–70. doi: 10.1016/j.jamda.2020.09.007

41. Chu CH, Donato-Woodger S, Dainton CJ. Competing crises: COVID-19 countermeasures and social isolation among older adults in long-term care. *J Adv Nurs*. (2020) 76:2456–9. doi: 10.1111/jan.14467