ORIGINAL ARTICLE

Basic and Translational Allergy Immunology

Allergy EUROPEAN JOURNAL OF ALLERGY AND CLINICAL IMMUNOLOGY EAACI | WILEY

Persistence of infectious SARS-CoV-2 particles for up to 37 days in patients with mild COVID-19

Tobias Zahn[1] | Ines Mhedhbi[1] | Sascha Hein[1] | Jan Raupach[1] | Csaba Miskey[2] |
Younes Husria[1] | Kathrin Bayanga[3] | Detlef Bartel[4] | Stefan Vieths[5] | Zoltan Ivics[2] |
Doris Oberle[6] | Brigitte Keller-Stanislawski[6] | Marie-Luise Herrlein[1] |
Thorsten Jürgen Maier[6] | Eberhard Hildt[1]

[1]Division of Virology, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

[2]Division of Medical Biotechnology, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

[3]Division of Transfusion Medicine, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

[4]Division of Allergology, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

[5]Management board, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

[6]Division of Pharmacovigilance, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, Germany

Correspondence
Thorsten Jürgen Maier, Division of Pharmacovigilance, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, D-63225 Langen, Germany.
Email: thorstenjuergen.maier@pei.de

Eberhard Hildt, Division of Virology, Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, D-63225 Langen, Germany
Email: eberhard.hildt@pei.de

Funding information
This research was funded by a grant b (VARIPATH) to EH from the Federal Ministry of Health (BMG)

Abstract

Background: People suffering from COVID-19 are typically considered non-infectious 14 days after diagnosis if symptoms have disappeared for at least 48 h. We describe three patients who independently acquired their infection. These three patients experienced mild COVID-19 and completely recovered symptomatically within 10 days, but remained PCR-positive in deep pharyngeal samples for at least 38 days. We attempted to isolate virus from pharyngeal swabs to investigate whether these patients still carried infectious virus.

Methods: Infectious virus was amplified in Vero E6 cells and characterized by electron microscopy and WGS. The immune response was investigated by ELISA and peptide arrays.

Results: In all three cases, infectious and replication-competent virus was isolated and amplified in Vero E6 cells. Virus replication was detected by RT-PCR and immunofluorescence microscopy. Electron microscopy confirmed the formation of intact SARS-CoV-2 particles. For a more detailed analysis, all three isolates were characterized by whole-genome sequencing (WGS). The sequence data revealed that the isolates belonged to the 20A or 20C clade, and two mutations in ORF8 were identified among other mutations that could be relevant for establishing a long-term infection. Characterization of the humoral immune response in comparison to patients that had fully recovered from mild COVID-19 revealed a lack of antibodies binding to sequential epitopes of the receptor-binding domain (RBD) for the long-term infected patients.

Conclusion: Thus, a small portion of COVID-19 patients displays long-term infectivity and termination of quarantine periods after 14 days, without PCR-based testing, should be reconsidered critically.
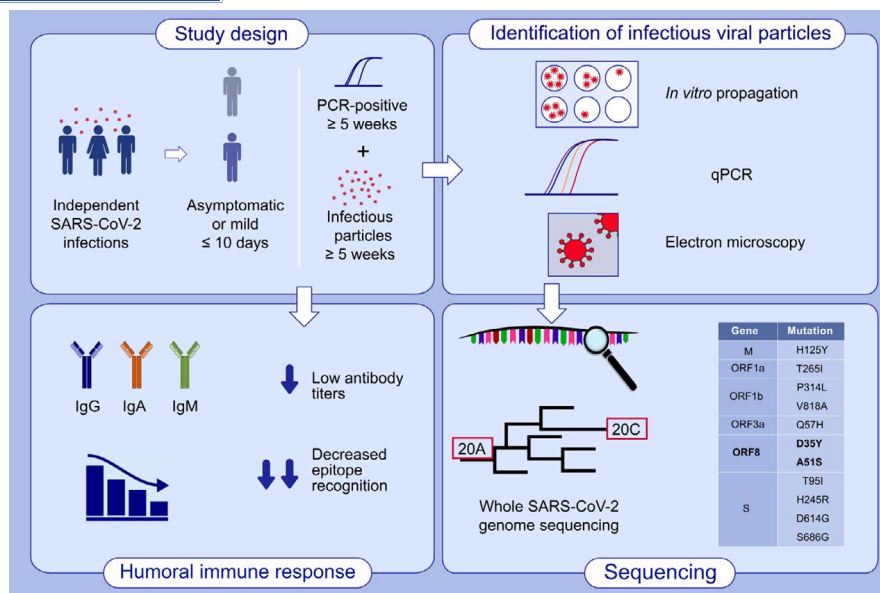
KEYWORDS
persistence, SARS-CoV-2, SARS-CoV2, mutations, whole-genome sequencing, antibodies

Thorsten Jürgen Maier and Eberhard Hildt contributed equally to this work.

Study design

Independent SARS-CoV-2 infections → Asymptomatic or mild ≤ 10 days

PCR-positive ≥ 5 weeks + Infectious particles ≥ 5 weeks

Identification of infectious viral particles

In vitro propagation

qPCR

Electron microscopy

Humoral immune response

IgG IgA IgM

Low antibody titers

Decreased epitope recognition

Sequencing

Whole SARS-CoV-2 genome sequencing

20A 20C

Gene	Mutation
M	H125Y
ORF1a	T265I
ORF1b	P314L
	V818A
ORF3a	Q57H
ORF8	**D35Y**
	A51S
S	T95I
	H245R
	D614G
	S686G

GRAPHICAL ABSTRACT

This study shows three long-term SARS-CoV-2 PCR-positive patients persistently producing infectious virus for up to 38 days without lasting symptoms. Patients displayed low antibody titers and decreased binding of IgG against sequential epitopes of spike protein in comparison to patients that fully recovered from mild COVID-19. WGS revealed that isolates belong to the 20A or 20C clade and viral eradication might be affected by two mutations in ORF8. Abbreviations: COVID-19, coronavirus disease 2019; M, membrane protein; ORF, open reading frame; qPCR, quantitative polymerase chain reactions; S, spike protein; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; WGS, whole genome sequencing.

1 | BACKGROUND

In December 2019, a novel lung disease was first reported to have emerged in Wuhan, Hubei Province, China, and showed very severe disease courses, including some with fatal outcomes. The new infectious disease rapidly spread throughout China and other countries. As the disease progressed, it was recognized to be a coronavirus infection, closely related to SARS (Severe Acute Respiratory Syndrome) and MERS (Middle East Respiratory Syndrome Coronavirus).[1-4] The disease has been referred to as COVID (Corona Virus Disease)-19 and the causative virus as SARS-CoV-2. The course of the disease can vary substantially in COVID-19 patients. Approximately 80 percent of detected infections are mild to moderate. In severe cases, respiratory distress and lung failure may occur. Severe courses mainly affect older people and individuals with underlying health conditions. However, severe cases can also be observed in people who are young or without pre-existing illness[5,6] The virus is primarily transmitted via droplet infection (ie, when speaking, sneezing, or coughing) and when there is little distance to other people. Transmission via aerosols is also possible.[7] At the time of writing this report, 225.778.167 total cases have been reported worldwide, with 4.648.145 deaths (source: COVID-19 Dashboard by the Center for Systems Science and Engineering [CSSE] at Johns Hopkins University [JHU]). Currently, contact restrictions, distance and hygiene rules, wearing a surgical mask in everyday life, testing for SARS-CoV-2 infection, and vaccination are the most powerful tools to keep infection events under control. There are reports describing that infectious virus could be readily isolated during the first week of symptoms from a considerable number of patients. To the best knowledge of the authors, so far no isolation of infectious virus was possible with samples of moderately or asymptomatically infected patients taken after day 10, although significant loads of viral genomes were detected.[8-11] Based on these findings, a common practice is to isolate COVID-19 patients for 14 days and to discontinue the quarantine without further testing given the patient is asymptomatic for at least 48 h. There are also reports that describe infectivity for at least 24 days after disease onset in patients with severe COVID-19 but showing no infectivity at later stages of a prolonged COVID-19 infection despite positivity in RNA testing.[12] Hence, because of the discrepancy of this finding to the experiments aimed at isolating active virus and the current time periods for quarantine, the timeline for discontinuation of transmission-based precautions is extensively debated.

This study aimed at addressing the question whether persistence of infectious virus particles for prolonged periods of more than 4 weeks is restricted to patients with severe COVID-19 or might also occur in patients with mild symptoms as well. In the present study, we show that in three patients who already had overcome a mild symptomatic COVID-19 disease active virus particles were still present in deep pharyngeal area for up to 37 days. Thus, there is an urgent medical need to understand the pathophysiology of the prolonged stay of active SARS-CoV-2 particles in some patients and to re-evaluate the procedures for discontinuation of transmission-based precautions of formerly infected COVID-19 patients.

2 | MATERIALS AND METHODS

2.1 | Cell culture and incubation with patients' swab

For this study, Vero E6 cells (ATCC® CRL-1586™) were used. Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM from BioWest) and supplemented with 2 mM L-glutamine, 100 µg/ml streptomycin, 100 U/ml penicillin, and 5% v/v fetal bovine serum (FBS superior; Sigma-Aldrich) under atmospheric conditions of 95% relative humidity and 5% CO_2 at 37°C.

Contents of SARS-CoV-2 positive or negative patients' swabs were resuspended in 1 ml of Vero E6 culture medium immediately after collection. The swab suspension was sterile-filtered through a 0.22 µm pore size sterile filter and added to one well of a 12-well cell culture plate with 1.5×10^5 Vero E6 cells per well. After 48 h, supernatant and cells were visually checked for indications of infections, such as color change of medium and especially microscopic examination of generated plaques, which reflect the cytopathic effect of SARS-CoV-2. Supernatant was removed and added to one well of a 6-well cell culture plate containing 3×10^5 Vero E6 cells per well. Additionally, 1 ml of Vero E6 cell culture medium was added to the well for a final volume of 2 ml. After 48 h (96 h after initial incubation), supernatant (2 ml) of one well of a 6-well plate was split into six wells of a 6-well plate containing 3×10^5 cells per well and filled with medium for a total volume of 1 ml per well. An additional 48 h later, supernatant was removed and inactivated by addition of TRIzol™ LS Reagent (ThermoFisher Scientific) for further sequencing analysis. Cells were washed with PBS and treated with TRIzol™ Reagent (ThermoFisher Scientific) for quantitative RT-PCR or washed with PBS and fixed with 3.7% formaldehyde for immunofluorescence staining and confocal laser scanning microscopy (CLSM).

2.2 | Titer quantification and plaque reduction assay

To determine viral titers in the supernatant of Vero E6 cells after inoculation with patients' swabs, supernatant after five passages was added as serial dilutions of 1:20–1:640 to Vero E6 cells. Cells were initially seeded as 2.5×10^5 cells per well in a 12-well plate and grown for 16 h at 37°C, 5% CO_2 and 90% humidity in DMEM supplemented with 2 mM L-glutamine, 100 µg/ml streptomycin, 100 U/ml penicillin, and 10% v/v FBS before inoculation. Vero E6 cells were inoculated for 1 h at 37°C in aforesaid medium without FBS. After that, medium was removed and 1 ml/well fresh medium containing 5% FBS was mixed with prewarmed liquid agarose (final concentration of 0.4% agarose) and added to the cells. Plates were left at room temperature for 20 min until agar got solid. Afterward, plates were incubated at 37°C for 72 h. For plaque visualization, 1 ml of 8% formaldehyde in PBS was added to each well and incubated for 20 min at 37°C. Agarose and formaldehyde were removed and cells were washed with PBS. Finally, 250 µl 0.1% crystal violet solution (Sigma-Aldrich) in 20% ethanol was added to the cells and incubated for 15 min at room temperature. Crystal violet solution was removed and cells were washed once with ddH_2O to quantify developed infection plaques.

For determination of neutralizing activity of sera of vaccinated individuals to isolates of long-term positive patients, sera dilutions of 1:20–1:640 were mixed with virus (final concentration 75 PFU/well). Virus was neutralized for 1 h at 37°C in medium without FBS. Afterward, the mixture was added to Vero E6 cells for 1 h at 37°C. An isolate of alpha-variant B.1.1.7 served as a control. Plaque quantification was performed as described above.

2.3 | RNA extraction and cDNA synthesis

RNA was extracted using Direct-zol™ RNA MiniPrep (Zymo Research). Reverse transcription of RNA to cDNA was performed by using M-MuLV Reverse Transcriptase RNase H negative (Genaxxon bioscience). RNA extraction and cDNA synthesis were performed according to manufacturer's protocols.

2.4 | Quantitative RT-PCR

Analysis of SARS-CoV-2 gene expression via q-RT-PCR was achieved by using a Maxima SYBR-Green qPCR Kit (Thermo-Scientific) and primers flanking SARS-CoV-2 N1 (nucleocapsid; FWD: 5′ GACCCCAAAATCAGCGAAAT-3′, REV: 5′-TCTGGTTACTGCCAGTTGAATCTG), N2 (FWD: 5′ TTAC AAACATTGGCCGCAAA-3′, REV: 5′-GCGCGACATTCCGAAGAA-3′), N3 (FWD: 5′ GGGAGCCTTGAATACACCAAAA-3′, REV: 5′-TGTA GCACGATTGCAGCATTG-3′), and RdRP (RNA-dependent RNA polymerase; FWD: 5 CAAGTGGGGTAAGGCTAGACTTT-3′, REV: 5′-ACTTAGGATAATCCCAACCCAT-3′) genes.[13] Measurements were performed with a LightCycler 480 Instrument II and analysis via LightCycler 480 Software 1.5.1.62 SP3 using absolute quantification fit points method (Roche).

2.5 | High-throughput sequencing library preparation

The total RNA isolated from the cell culture supernatant was used for Illumina library preparation using a modified NNSR priming method.[14] We removed the rRNA from the samples with a QIAseq FastSelect- rRNA HMR kit (Qiagen) using the recommended reaction with the following program: 2 min, 75°C; 2 min, 70°C; 2 min, 65°C; 2 min, 60°C; 2 min, 55°C; 5 min, 37°C; 5 min, 25°C. The above reaction was combined with subsequent reverse transcription with 100 pmol NNSR_RT primer, 10 mM DTT, 0.05 mM dNTPs, 20 Us of RiboLock RNase inhibitor, 4 µg Actinomycin D, and 200 Us Superscript IV (Thermo Fisher Scientific) as follows: 45°C for 30 min; 70°C for 15 min. The cDNA was purified with 1.8 volumes

of magnetic beads, and RNase H digest was performed as recommended (NEB). After repeated bead purification, the cDNA samples were subjected to second-strand synthesis at 37°C for 30 min in a 50 μl final reaction volume containing NEB buffer 2, 0.125 mM dNTPs, 5 Us of 3′-5′exo-Klenow Fragment (NEB), and 200 pmol NNSR-2 primer. The resulting double-stranded DNA was PCR-amplified to obtain the barcoded Illumina libraries with a NEBNext Ultra II Master Mix (NEB), containing 25 pmol NNSRnest_ind_N and NNSR_Illumina primers, each, using the following cycling conditions: 98°C 10 s; 5 cycles of 98°C 10 s, 55°C 30 s, 68°C 30 s; 15 cycles of 98°C 10 s, 65°C 30 s, 68°C 30 s. The resulting 400–700-bp-long DNA smears were isolated from 1.5% agarose gels with a Zymoclean Gel DNA Recovery Kit (Zymo Research), quantified with qPCR using a NEBNext Library Quant Kit for Illumina (NEB), and sequenced on a MiSeq instrument with a paired-end 2×301 setting.

The overall quality of the reads were checked with the FastQC tool.[15] Raw reads were trimmed and aligned against the reference sequence of SARS-COV-2 Wuhan-Hu-1 isolate (Accession number: NC_045512) with Trimmomatic (v0.39) and SNAP, respectively.[16,17] After sorting, the consensus call was performed with Rsamtools and SAMtools.[18,19] The variant and phylogenetic analysis was performed with Nextclade (https://clades.nextstrain.org/).[20] The phylogenetic tree was visualized with auspice, which is part of the Nextstrain project.[20]

2.6 | Virus concentration, purification, and transmission electron microscopy

After inoculation of Vero E6 cells with swabs of SARS-CoV-2 PCR-positive and -negative patients and viral amplification, 5 ml of virus-containing cell culture supernatant was inactivated overnight by addition of formaldehyde to a final concentration of 4%. Virus concentration and purification was performed according to.[21]

For microscopic, negative stain visualization of viral particles, transmission electron microscopy (TEM) was used. A drop of 20 μl of sample was pipetted on a carbon-coated, glam-discharged formvar grid and immobilized for 30 min at RT. After two washing steps with 25 μl ddH2O each, grids were incubated with 1% phosphotungstic acid in ddH$_2$O for 10–30 s at RT. Grids were dried and analyzed using an EM 109 (Zeiss).

2.7 | ELISA

Receptor-binding domain protein was produced in HEK293T cells and purified with Ni-NTA affinity chromatography according to the protocol from.[22] Florian Krammer, University of Icahn School of Medicine at Mount Sinai, kindly provided the expression plasmid pCAGGS-sRBD. 50 μl at a concentration of 4 μg/ml RBD in PBS was used to coat 96-well microtiter plates (costar 3590, Corning Incorporated) overnight at 4°C. Afterward, plates were blocked for 1 h at RT with 10% FCS in PBS. Between each step, plates were washed three times with PBS-T (PBS supplemented with 0.05%

Tween). The sera were prediluted 1:100 in 10% FCS in PBS and incubated for 1.5 h at RT. For detection, anti-human IgG HRP-linked ECL antibodies (Cytiva) were used in a 1:3000 dilution. After 1 h of incubation at RT, ELISA plates were developed with 75 μl TMB ELISA Substrate Solution (eBioscience) for 5 min, stopped with 75 μl 1 N sulfuric acid, and analyzed directly at 450 nm on a Tecan reader (Tecan Group). As a negative control, sera of patients without SARS-CoV-2 infection nor vaccination, collected 2017, were used. All values are given as absorbance at 450 nm. The dotted line, representing positivity cut-off, was determined in respect of negative control.

2.8 | Peptide array

For mapping of sequential stretches of the SARS-CoV-2 proteins full-length spike (S), nucleocapsid (N), membrane (M), and envelope (E) protein recognized by antibodies in sera of SARS-CoV-2 PCR-positive and -negative patients were used. In approximation of epitopes, 253 (S), 52 (N), 27 (M), or 9 (E) synthetic overlapping peptides of 15 aa in length with an offset of 5 (S) or 8 (N, M, E) aa between each peptide, representing the respective proteins' sequences, were synthesized according to the principle of Merrfield using a MultiPep RSi automated peptide synthesizer (Intavis AG). Using a slide-spotting robot (Intavis AG) peptides were immobilized in a duplicate on a cellulose membrane (Intavis AG) attached to a 76 mm × 26 mm slide.

Slides were pre-incubated with 10% 10× Casein Blocking Buffer in TBS-T, pH 7.4, supplemented with 5% sucrose for blocking purposes and incubated with patients' sera, 1:40 diluted in casein blocking buffer. For antibody detection, slides were incubated with LI-COR antibody anti-human-IgG-800 secondary antibody (1:10.000 diluted in 1:10 Casein Blocking Buffer) and bound antibodies were detected via an LI-COR Odyssey Infrared Imager (Biosciences). The peptide array was validated with sera of convalescent individuals. Sera of these persons were diluted to comparable ODs of the lowest measured IgG OD of the long-term infected ones (~OD 0.4) and compared directly to each other in an individual approach. The results of these peptide arrays showed much stronger binding to defined epitopes in case of the adjusted convalescent sera as compared to the sera derived from the long-term infected patients.

2.9 | Immunofluorescence staining and CLSM analysis

Incubated cells were fixed with 3.7% formaldehyde and permeabilized with 0.5% Triton X-100 in PBS. Afterward, cells were blocked with 1% (w/v) bovine serum albumin in TBS-T (20 mM Tris, 150 mM sodium chloride pH 8.8, 0.05% [v/v] Tween 20) and incubated with primary antibody rabbit-anti-SARS-CoV-2 spike protein (40150-T62-COV2-100, Sino Biological; diluted 1:800 in PBS). For detection of primary antibodies, Cy3-coupled donkey-anti-rabbit antibody (Jackson ImmunoResearch Laboratories, Inc; diluted 1:400 in PBS) was used. Nuclei were stained with 4′,6-Diamidin-2-phenylindole (DAPI; diluted

1:1000 in PBS). CLSM analysis was performed using a Leica SP8 confocal laser scanning microscope (Leica).

2.10 | Ethics statement

The study (PEI-SARS-CoV2) was approved by the local ethics committee (Ethik-Kommission, Landesärztekammer Hessen 60314 Frankfurt am Main) and written informed consent was obtained from all patients (2020-1664_2-evBO). Patients provided a written agreement to publish their medical data. The study was performed in accordance with the provisions of the Declaration of Helsinki and good clinical practice guidelines.

3 | RESULTS

3.1 | Identification of long-term infected COVID-19 patients

A physician-led medical service (Internal Medical Service, IMS) was established at the Paul Ehrlich Institute (PEI) in Langen/Hessen/Germany in April 2020 to conduct routine and occasion-based testing of employees and to provide counseling and care for employees

and families affected by COVID-19. As part of the testing, personal information and possible symptoms are also collected with the consent of the person being tested. Formerly infected employees can only return to the PEI after being tested negative for SARS-CoV-2 using PCR run on samples obtained from throat swabs. At the time of writing, 26 cases of COVID-19 were registered among employees at the PEI. Three infected employees (patients 1–3) were found to still be positive for SARS-CoV-2 at day 37, 33 and 31, respectively, after the initial PCR-based diagnosis. The local health agencies had already released all three employees without any restrictions from home quarantine 14 days after the initial diagnosis. An overview of the infection periods of all long-term infected patients and sample taking is given in the time line (Figure 1).

3.2 | Patient´s medical history and anamnesis

Patient 1 is male and 27 years old. The patient was tested positive on pharyngeal SARS-CoV-2-RNA using RT-PCR with $C_t = 16$ on 11/30/2020 by a general practitioner at the Institute for Medical Diagnostics Bioscientia GmbH. Symptoms with tightness in the chest and increased temperature of up to 38°C had already existed 2 days earlier since 11/29/2020. The symptoms lasted for 11 days and mainly consisted of tightness in the chest, cough, headache, loss

FIGURE 1 Overview of the infection periods of all long-term infected patients and of sample taking Overview about sampling times and performed tests

of smell and taste, and toward the end of the illness, increased tiredness. Apart from a known hypothyroidism since 2010 under substitution with iodide 200 μg once a day, the patient reported no further diseases. The patient did not take any antivirals, immunosuppressive drugs or other medication before or during his infection. The last collection of material from a deep pharyngeal swab (IMS, PEI) and positive reactivity ($C_t = 35$) in the RT-PCR for SARS-CoV-2 was conducted on 01/06/2021 (Bioscientia). The PCR assay performed on 01/08/2021 already showed a negative reactivity. Thus, the patient had been documentably infected with SARS-CoV-2 for at least 37 days. Considering the initial symptomatic stage and a possible SARS-CoV-2-positivity even after 01/07/2021, the probable duration of the infection was approx. 39–40 days. On 01/08/2021, a swab was collected for in vitro replication of SARS-CoV-2 in cell culture.

Patient 2 is male and 36 years old. The patient was tested positive on pharyngeal SARS-CoV-2-RNA by a general practitioner on 12/07/2020 by RT-PCR (MVZ Dr. Helge Riegel GmbH). C_t value was 28.06. Prior to diagnosis, there was already symptomatology with sinusitis, headache, and increased temperature up to 38°C since 12/01/2020. Symptoms lasted over 14 days and consisted of nausea, loss of smell and taste, muscle pain, and increased fatigue, primarily at the end of the illness period. Apart from migraine, the patient reported no other underlying diseases. No antivirals, immunosuppressive drugs or other medications were administered to this patient before or during his infection. The last collection of material from a deep pharyngeal swab taken by the IMS with positive reactivity at $C_t = 31$ using RT-PCR for SARS-CoV-2 (Bioscientia) was performed on 01/08/2021. The PCR analysis on 01/11/2021 (Bioscientia) demonstrated a negative reactivity. Thus, the patient had been documentably infected with SARS-CoV-2 for at least 33 days. Including the initial symptomatic stage and a possible SARS-CoV-2-positivity later than 01/08/2021, the probable duration of infection was, therefore, approximately 39–41 days. A swab was also collected for in vitro replication of SARS-CoV-2 in cell culture on 01/08/2021.

Patient 3 is female and 36 years old. The patient was tested positive on SARS-CoV-2-RNA by RT-PCR by the IMS at the PEI on 12/15/2020 (Division of Virology, PEI) during routine testing. A validation of the positive reactivity was obtained by an external diagnostic laboratory (sampling: IMS, analysis: Bioscientia) on 12/16/2020 using PCR with a C_t value of 15. Interestingly, the patient did not develop any symptoms at any time that would have indicated an infection with SARS-CoV-2. The patient reported a lack of underlying diseases. The patient did not take any antivirals, immunosuppressive drugs or other medication before or during her infection. The last collection of material from a deep pharyngeal swab with clearly positive reactivity using RT-PCR for SARS-CoV-2 was conducted on 01/14/2021 (IMS) with $C_t = 31$. On 01/18/2021, there was a reported ambiguous PCR reactivity to SARS-CoV-2, and on 01/21/2021, only an initial weak reactivity was observed during the PCR run, which was evaluated as questionably negative. The PCR analysis conducted on 01/21/2021 then resulted in a clear negative

reactivity. Thus, the patient had been documentably infected with SARS-CoV-2 for at least 31 days. Including the possibility of an infection before 12/15/2020 and the unclear terminal reactivity in PCR analysis that was observed until 21/01/2021, the probable duration of infection was approximately 35–40 days. A swab was also collected for in vitro replication of SARS-CoV-2 in cell culture on 01/18/2021.

As control, three documentably SARS-Cov-2-negative, healthy volunteers P4-6 were included in this study and material from swabs was subjected to in vitro replication following the identical procedure as described for patients 1–3.

3.3 | Detection of infectious SARS-CoV-2 in pharyngeal swabs of long-term COVID-19 patients

To determine whether the presence of viral nucleic acids in pharyngeal swabs of patients (P1–P3) who had undergone COVID-19 weeks before, recovered, and lacked any symptoms was based on the presence of infectious virus, we attempted to amplify virus from pharyngeal swabs. For this purpose, pharyngeal swabs were dispensed in medium, sterile-filtered, and added to Vero E6 cells. Pharyngeal swabs from PCR-negative patients served as controls (P4–P6). To exclude any signals that may have been based on input material, virus amplification was performed by two passages over 5 days or over 7 days with five passages. Titer determination via plaque assay after five passages revealed titers of 5.46×10^4 PFU/ml for P1, 5.6×10^4 PFU/ml for P2, and 5.89×10^4 PFU/ml for P3. To detect viral replication, RNA was isolated from cellular lysates and from cell culture supernatant at both time points. After cDNA synthesis, SARS-CoV-2-specific RT-PCR was performed utilizing three gene segments of the nucleocapsid (N)-gene and one of the RNA-dependent RNA polymerase (RdRP) gene. RT-PCR analysis revealed in the case of isolates from P1 low Ct values (approximately 10 for the intracellular RNA and 18 for the extracellular RNA) after 5 days (Figure 2A,B). For P2 and P3, Ct values of about 10 were observed after 7 days for intracellular RNA and roughly 23 for extracellular RNA. In contrast, all negative controls showed high values of approximately 40, comparable to the water control. For all positive samples, detection of the RdRP gene always showed the highest Ct values in all patients. Because the positive RT-PCR data were obtained after several rounds of passaging and reflect a significantly higher genome number as compared to the input, contamination by input RNA can be excluded.

To corroborate these data that indicate the presence of infectious replication-competent virus in the pharynx-derived material, Vero E6 cells, inoculated with the pharynx-derived material for 7 days and four passages, were analyzed by immunofluorescence microscopy (CLSM) using spike-specific antibodies (Figure 3). The immunofluorescence microscopy revealed viral replication evidenced by the significant production of spike proteins in the case of cells infected with the isolates from P1 and P2. For P3, only some cells showed spike-specific signals.

FIGURE 2 Analysis of viral RNA after inoculation of Vero E6 cells with patients' swabs. Vero E6 cells were inoculated with contents of swabs of either SARS-CoV-2 long-term PCR-positive patients (P1-P3) or PCR-negative volunteers (P4-P6). Supernatants were passaged for 5 days with a total of total passages or for 7 days with a total of three passages. For detection of SARS-CoV-2 specific signals, primers flanking N1, N2, N3 and RdRP genes of SARS-CoV-2 were used. RNA was isolated from Vero E6 cell lysates (A) or from culture supernatant (B). Vero E6 cells without infection served as a negative control, cells infected with SARS-CoV-2/2020/FR/702 served as a positive control
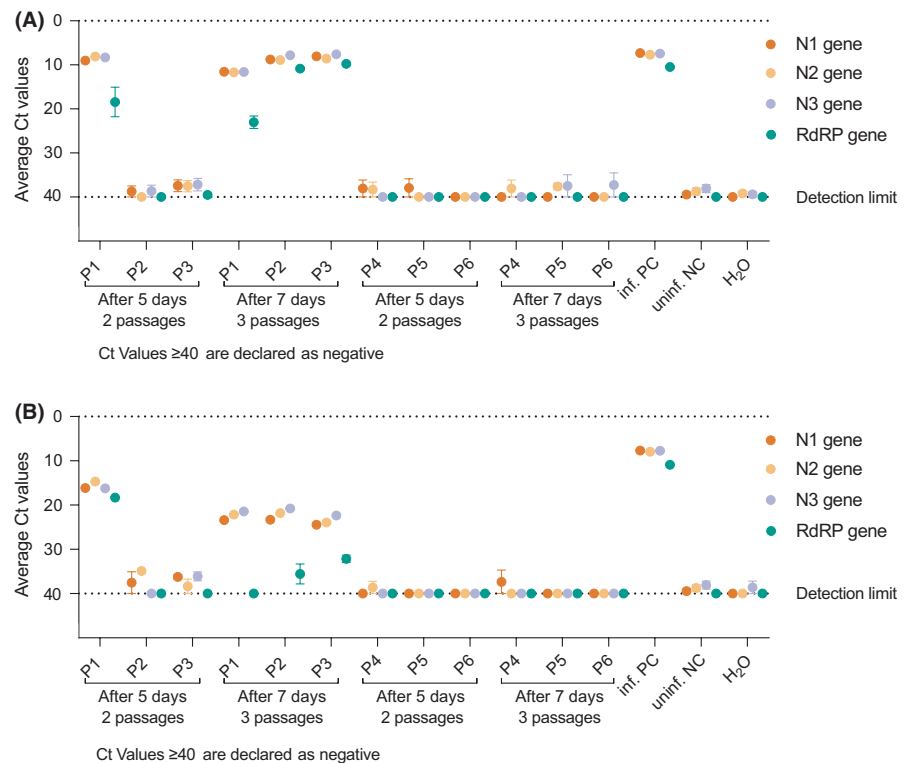

FIGURE 3 Immunofluorescence microscopy of infected Vero E6 cells. Vero E6 cells, inoculated with patients' swabs of either long-term PCR-positive patients (P1-P3) or −negative volunteers (P4-P6), were analyzed via immunofluorescence microscopy. Cells were stained with anti-SARS-CoV-2 spike antibodies (red) for detection of virus and with DAPI (blue) to visualize nuclei. Vero E6 cells without infection served as a negative control, cells infected with SARS-CoV-2/2020/FR/702 served as a positive control
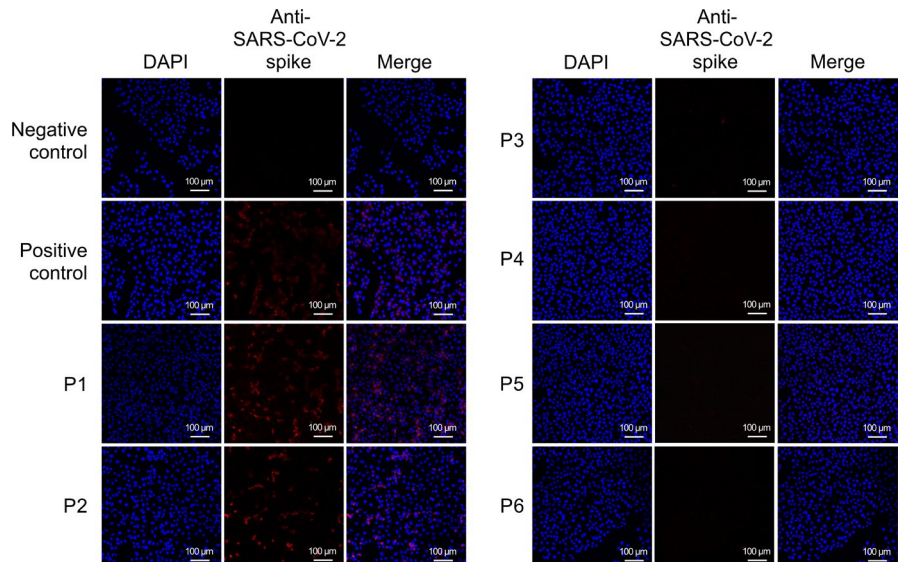
Taken together, these data indicate the presence of replication-competent infectious virus in the pharyngeal tract of all tested long-term PCR-positive patients.

3.4 | Detection of coronavirus particles by electron microscopy

To further confirm the formation of intact viral particles, virions were isolated by ultracentrifugation and analyzed by TEM after negative staining. The electron microscopy clearly showed the presence of viral particles with the characteristic morphology of coronaviruses in the amplified isolates of all three tested patients (Figure 4). The size of viral particles, with a diameter of approximately 75 nm, the length of spike structures on the surface of these particles of about 15 nm, and the general morphology corresponded to SARS-CoV-2 particles (Figure 4D). Taken together, the electron microscopy confirmed the findings of the RT-PCR and immunofluorescence microscopy: the pharyngeal swabs obtained from patients 1, 2, and 3 contained infectious, replication-competent SARS-CoV-2.

3.5 │ SARS-CoV-2 isolated from long-term infected patients belonged to 20A and 20C clade

To characterize the isolates in more detail and classify the isolates, Next-Generation Sequencing (NGS) analysis was performed. For this purpose, RNA was isolated after a single passage to minimize the risk of in vitro mutations. Based on the sequence analysis, the isolates derived from P1 and P3 could be identified as belonging to the 20C clade, whereas that of P2 belongs to the 20A clade (Figure 5A). (GenBank accession numbers; Patient 1: OK075089, Patient 2: OK075090, Patient 3: OK075091).

The detailed analysis of the whole-genome revealed eleven mutations in P1 and P3 and ten mutations in P2 compared to the aligned strain NC_045512.2 (Table 1). These mutations include the common mutations D614G of the spike protein as well as the often co-occurring mutations P314L in ORF1b/RdRP and T265I in ORF1a.[23-26] Additionally, the also already well-known Q57H mutation in ORF3a

(not detected in isolate of P2) and the S686G mutation in the spike that was originally described in virus isolates of ferrets were identified.[24,27] Besides, the more uncommon mutations H125Y in the M protein, the mutations V818A in ORF1b, D35Y and A51S in ORF8, and T95I and H245R in the spike protein could be identified. No particular abnormalities could be detected within the receptor-binding domain (RBD).

3.6 │ Lack of antibodies binding to sequential epitopes in the RBD

To study the humoral immune response of the long-term infected patients, the amount of SARS-CoV-2 RBD antibodies were determined by ELISA and stratified for IgG, IgA, and IgM. Sera were collected at day 38 for patient 1, after 36 days for patient 2, and after 31 days for patient 3 after initial PCR analysis. For all three patients,
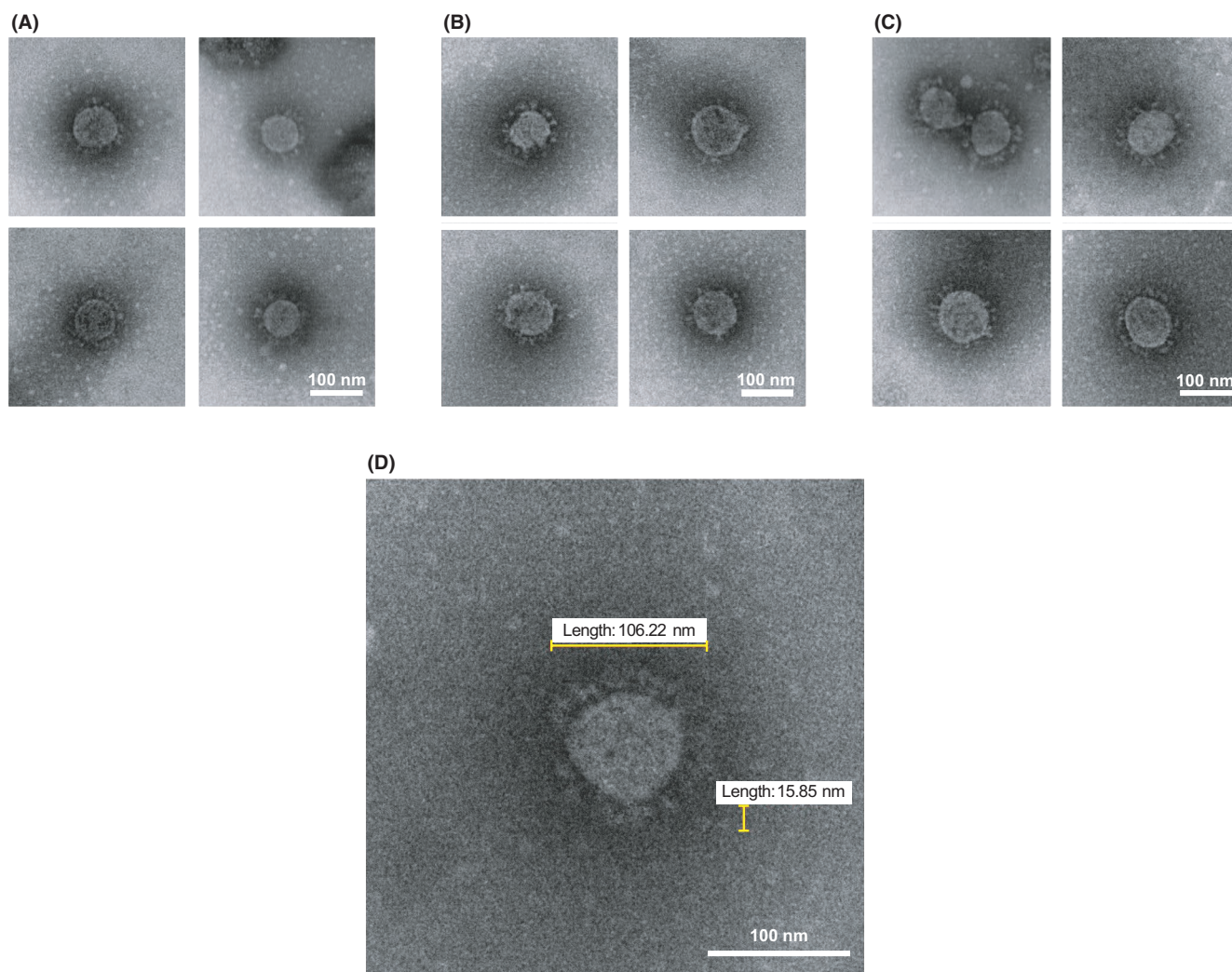

FIGURE 4 Visualization of isolated SARS-CoV-2 particles via TEM. Supernatant of infected Vero E6 cells was collected and virus was concentrated for microscopic, negative stain visualization of particles. Transmission electron microscopy (TEM) showed the presence of SARS-CoV-2 particles in the supernatant of Vero E6 cultures infected with swabs of P1 (A), P2 (B) and P3 (C). The analyzed particles are showing the typical size and morphology of SARS-CoV-2 (D). Scale bar is 100 nm
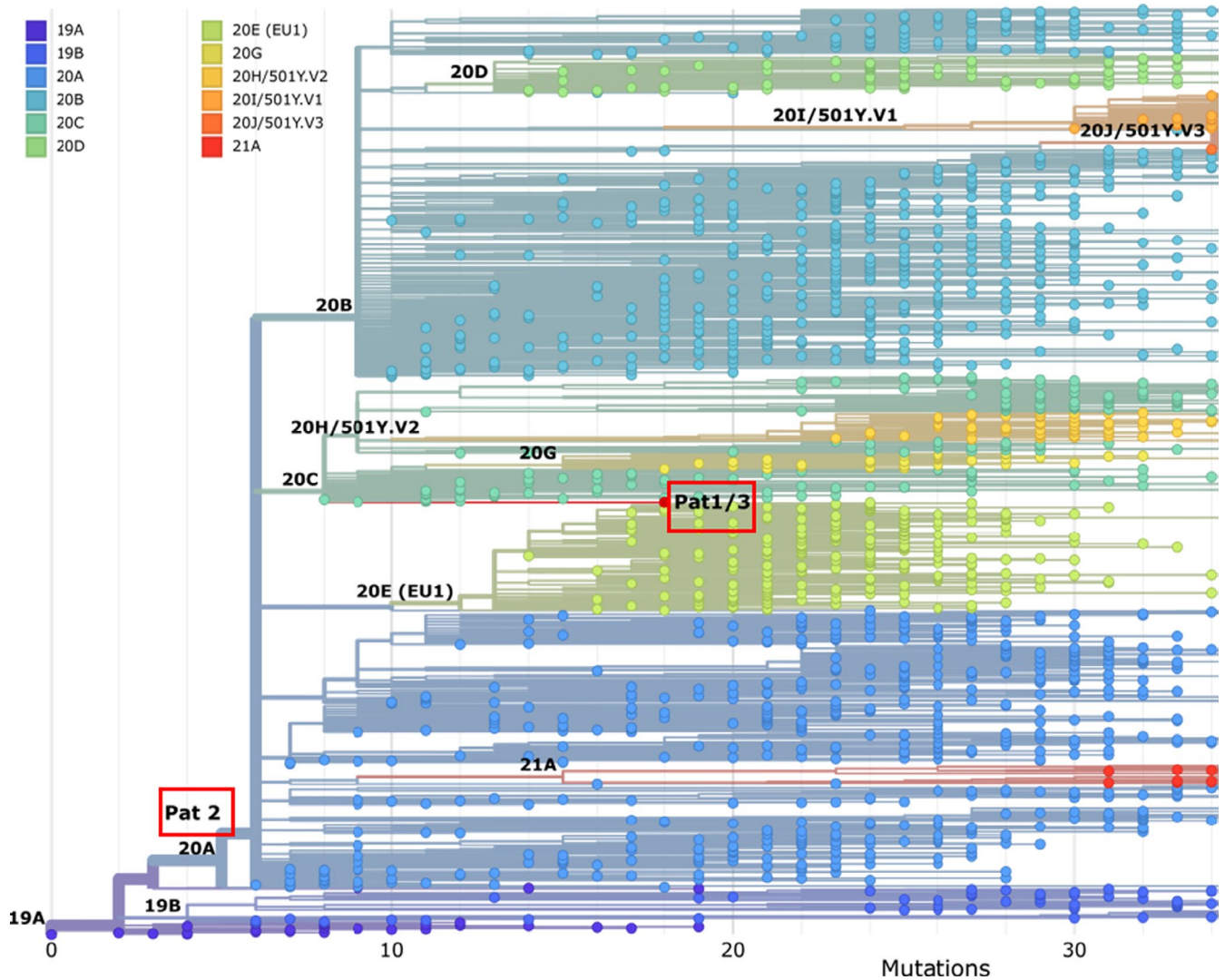
FIGURE 5 Phylogenetic characterization of isolated RNA of infected Vero E6 cells. Sequence analysis of RNA from supernatant of swab inoculated Vero E6 cells identified the presence of SARS-CoV-2 RNA, belonging to Clade 20A or 20C

TABLE 1 Identified mutations after whole-genome sequencing

Gene	Mutation
M	H125Y
ORF1a	T265I
ORF1b	P314L
	V818A
ORF3a	Q57H (not in P2)
ORF8	D35Y
	A51S
S	T95I
	H245R
	D614G
	S686G

the presence of RBD-specific antibodies was detected (Figure 6). For patient 1, a higher IgG response was found as compared to patients 2 and 3. The highest IgM response was found for patient 2, whereas for patients 1 and 3, a moderate level was detected. A low IgA level was found for patients 1 and 2, and for patient 3, the level was moderate. Compared to convalescent sera of SARS-CoV-2 infected subjects, IgG and IgA antibody levels of all long-term PCR-positive individuals showed lower values, but higher IgM levels.

For a more detailed analysis of the antibody response, peptide arrays covering the S, N, M, and E protein were established, and the binding pattern of the sera derived from patients 1–3, from three negative patients, and from three convalescent plasma samples were determined (Figure 7). It should be emphasized that this technique mainly reflects binding to sequential epitopes. To identify specifically recognized peptides, the signals obtained for sera from SARS-CoV-2 negative samples were considered as background (Figure 7A).

If the background signals were subtracted from the specific signals (Figure 7C), there was only one sequential epitope specifically detected in the N protein by serum derived from patient 1. In case of the serum derived from patient 2 two linear epitopes in the S1 domain of the spike, one epitope derived from the N protein
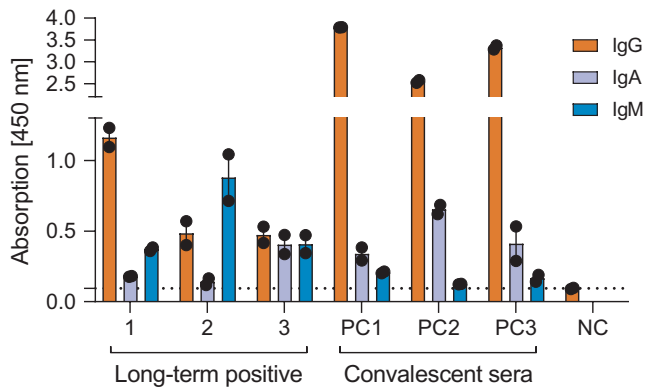
FIGURE 6 Antibody levels of long-term SARS-CoV-2 PCR-positive patients. Anti-SARS-CoV-2 antibodies of long-term positive patients were measured via SARS-CoV-2 RBD-specific ELISA. Thereby, humoral response was differed in IgG, IgA and IgM levels. Collection of sera for antibody analysis were performed 38 days (P1), 36 days (P2) and 31 days (P3) after initial PCR testing. SARS-CoV-2-negative, nonvaccinated patients' sera served as negative control. As a positive control and comparison, three convalescent sera of infected individuals (PC1-3) without long-term PCR positivity were used

and one from the M protein were identified. In contrast to this, the linear epitopes recognized by serum derived from patient 3 were completely different from the patterns observed for the other patients. Here, we identified one epitope at the C-terminal part of the spike S2 domain and one in the N protein (Figure 7A,C; Table 2). In contrast, a wide variety of sequential epitopes was recognized with high intensity by convalescent sera (Figure 7B) derived from patients who rapidly eliminated the virus (Figure 7B). These epitopes were not recognized by sera derived from long-term patients (Table 3). In the case of the convalescent sera, epitopes localized within the RBD were recognized. This finding is in clear contrast to the binding pattern obtained for the long-term patients; here, no specific binding to linear epitopes covering the RBD was observed. Taken together, these data indicate that the pattern of sequential epitopes differs for sera derived from long-term infected patients and from patients that eliminated the infection within 10 days.

3.7 | Reduced neutralization of isolate P1 by sera of vaccinated individuals

To determine how efficiently infection with isolates of long-term positive patients can be neutralized by sera of Comirnaty (BNT162b2, Biontech/Pfizer) vaccinated individuals, a plaque reduction neutralization test (PRNT) was performed and compared to neutralization of an isolate of the SARS-CoV-2 alpha-variant (B.1.1.7; Figure 8). First, neutralization tests showed different neutralizing capacities of the three used sera, with low (S1), medium (S2) und high (S3) neutralizing capacity concerning all isolates. While isolates P2 and P3 could be neutralized by S1 at dilutions lower then 1:320, this serum blocked infection with the alpha-variant and P1 at 1:80 and ≤1:80. Medium

neutralizing serum S2 could block infection with alpha at a dilution of about 1:320 and P2 and P3 at a dilution of ≤1:640, and a dilution of 1:40 was needed to neutralize infection with P1 isolate. While S3 neutralized infection with isolates alpha, P2 and P3 in dilutions of 1:640, it shows slightly less efficiency neutralizing P1 with ≤1:640. These data suggest that isolates of P2 and P3 are more efficiently neutralized by sera of vaccinated individuals compared to SARS-CoV-2 variant B.1.1.7. However, infection with isolate of P1 could be less efficiently blocked by all sera compared to the other isolates of long-term positive patients or to the alpha-variant.

4 | DISCUSSION

The data presented in this study demonstrate that in non-immunocompromised patients, SARS-CoV-2 infection may persist for several weeks. Moreover, these results indicate that the detection of viral nucleic acids even weeks after diagnosis of infection does not only reflect remnants of viral genomes. There were still infectious viral particles detectable in the pharynx as demonstrated by amplification and detection the intact and functional virus. Apart from the RT-PCR data, immunofluorescence microscopy, WGS, and electron microscopy unequivocally showed that infectious virus was present in the pharyngeal swabs of the patients, which could be isolated and amplified. Because the quantification of infectious viral particles depends on the quality of the swab, conclusions on the viral load in the respective patients require a critical assessment. However, in all three patients, even more than 4 weeks after COVID-19 diagnosis the positive RT-PCR results obviously correlated with the presence of infectious viral particles, reflecting an ongoing active infection.

A detailed analysis by NGS revealed that SARS-CoV-2 isolated from P1 and P3 belonged to clade 20C, whereas isolate of P2 belonged to clade 20A and sequence analysis revealed several mutations. The 20A clade first attracted attention in beginning of 2020 with increasing outbreaks in Europe, followed by displacement of other lineages in North and South America.[28] It is particularly characterized by its D614G mutation in the spike protein, caused by A23403G nucleotide change.[29] Clade 20C on the other hand represents a daughter clade of 20A, also harboring the D614G mutation.[30] This mutation had already been described as part of a widespread variation, leading to increased infectivity due to facilitated ACE2-binding characteristics.[23,24] This D614G mutation is known to be frequently accompanied by the P314L in ORF1b/RdRP that had also been identified in the patient 1 isolate (Table 1).[25] Another common mutation that could be detected in these isolates was T265I in ORF1a. The threonine substitution by isoleucine at position 265 leads to the addition of a beta-sheet structure.[26] We could identify the Q57H mutation in ORF3a in the isolates of P1 and P3, which was first described in Singapore in February 2020. This mutation coexists in many cases with the mutation T85I in non-structural protein 2 (NSP2) that, in contrast, could not be identified in this isolate.[24] One mutation that was originally described in ferrets and was
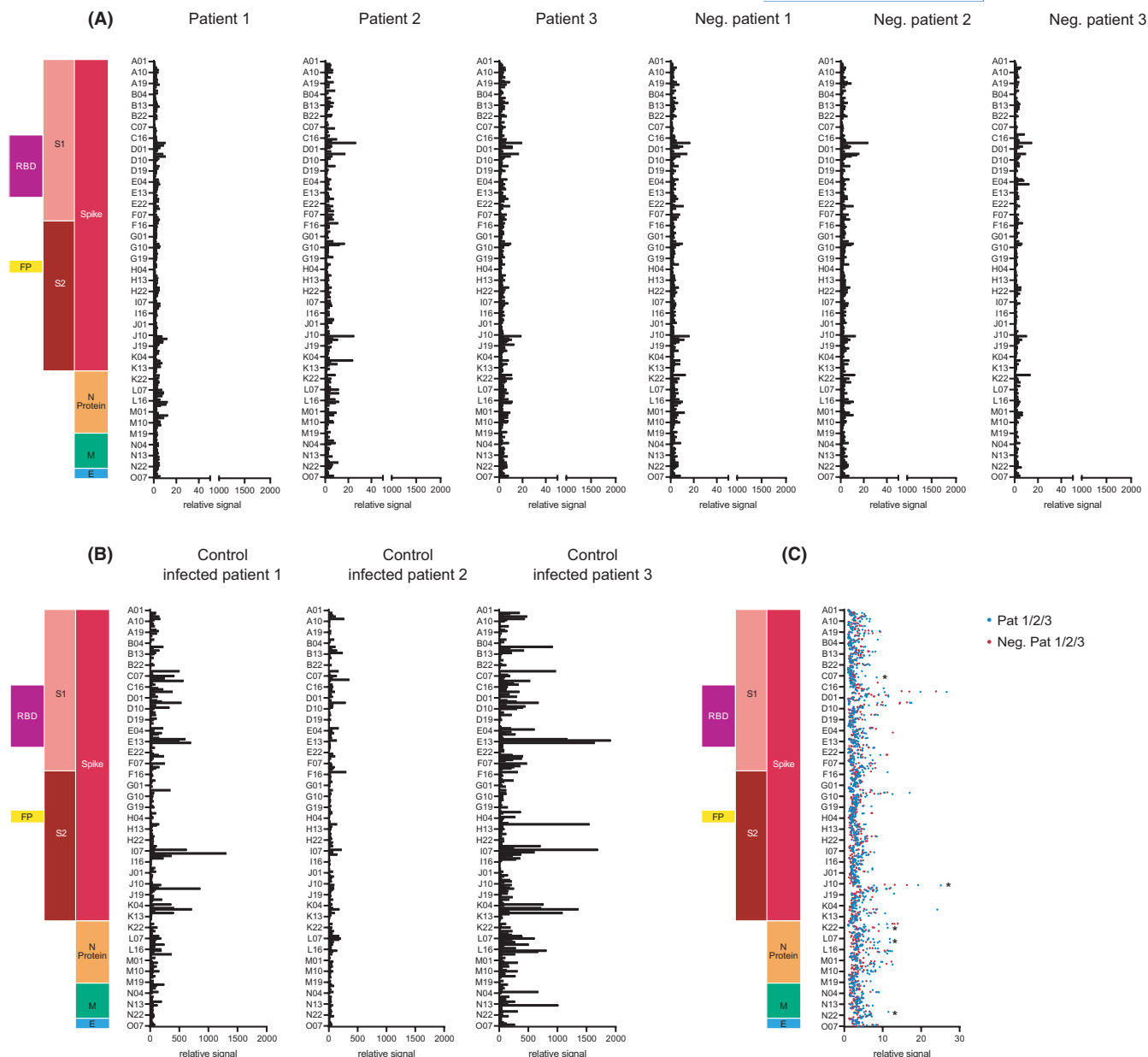
FIGURE 7 Epitope mapping of patient derived antibodies against SARS-CoV-2 peptides. 253 (S), 52 (N), 27 (M), or 9 (E) synthetic overlapping peptides (15 aa in length; 5 (S) or 8 (N, M, E) aa offset between each peptide) were spotted as duplicates onto microscope slides and incubated with sera of long-term PCR-positive and –negative patients (A). Sera from convalescent patients who were not long-term PCR-positive served as further control (B). Putative sequential epitopes were determined by identifying peak signals of both positive and negative patients, with eliminating possible epitopes detected by sera of positive patients if signals were identified even to a minor amount in negative patients, too. Selectively recognized specific epitopes were marked by a star (C)

supposed to lead to immune evasion before receptor binding was the S686G mutation in the spike protein that we could also identify as part of the analyzed sequence. A rare missense mutation that was detected in our study is H125Y in the M protein. It is not clear what impact that mutation might have in the context of stability or immune evasion, but it seems to be a mutation with globally increasing incidence.[31] Furthermore, next to the well-recognized mutations D614G and P314L in ORF1b a new and so far unknown mutation V818A in ORF1b, was detected. Because ORF8 is known to play a crucial role in inhibiting the type I interferon signaling pathway as

well as major histocompatibility complex I (MHC I) degradation, it was interesting to detect two mutations in ORF8.[32,33] As MHC I degradation facilitates immune evasion, it was an unexpected finding that mutation D35Y that we identified in ORF8 was formerly described as providing protein stability and decreasing disease severity.[34,35] The second ORF8 mutation A51S, in contrast, is located within two possible CD4+ T cell epitopes.[36] The consequences arising from this mutation had yet not been investigated. However, these epitope mutations could possibly affect the T cell response against SARS-CoV-2 positive cells and delay virus elimination by the

cellular immune response. In addition to these mutations affecting T cell epitopes, two further mutations in the spike protein were identified that could affect the T cell response. This includes T95I, a mutation occurring within a predicted CD8+ T cell epitope,[37] and mutation H245R that might be located both within a T cell and/or B cell epitope.[38] Therefore, it can be hypothesized that these mutations contribute to impaired virus elimination, as reflected by a productive long-term SARS-CoV-2 infection.

Moreover, we found distinct differences in relative antibody concentrations in serum between the three long-term patients as compared to reconvalescent patients. In the majority of SARS-CoV-2 patients, IgG and IgM is detectable 2–3 weeks after onset of symptoms.[39] For patients with no obvious emerging symptoms or who remained completely asymptomatic, IgG and IgM levels are consistently less pronounced compared to patients with noticeable mild, moderate, or severe symptoms.[40] Patients 1 and 2 in this study both showed SARS-CoV-2 infections with similar symptoms, but only patient 1 demonstrated an IgG seroconversion that resulted in a considerable increase of IgG antibody titers (Figure 6), whereas IgG levels in patient 2 were similar to asymptomatic patient 3. Another striking feature of the antibody response in patient 2 are the levels of IgM clearly exceeding those of IgG after 38 days of symptom onset, albeit IgG levels are typically higher than IgM after such time period.[41] Furthermore, symptomatically infected patients 1 and 2 surprisingly produced only a rather weak IgA response at all, even though other studies described early and rather high IgA titers beginning from approximately 10 days to more than 40 days and even longer in symptomatically infected patients.[42,43] The lowest IgA titers were observed in both symptomatically infected patients 1 and 2, and in both cases, swab inoculation led to distinct in vitro infections (Figure 3A), whereas the highest IgA titers were measured in asymptomatic patient 3. Here swab inoculation caused only weaker infection in vitro (Figure 3B). These observations strengthen the hypothesis of a correlation of IgA titers with the efficacy of virus neutralization.

TABLE 2 Sequential epitopes identified in infected long-term patients

Patient 1	
Epitopes: (identified outliers)	
N	G-A-L-N-T-P-K-D-H-I-G-T-R-N-P
Patient 2	
Epitopes: (identified outliers)	
S1	S-L-L-I-V-N-N-A-T-N-V-V-I-K-V
	Q-P-R-T-F-L-L-K-Y-N-E-N-G-T-I
N	L-P-Y-G-A-N-K-D-G-I-I-W-V-A-T
	H-I-G-T-R-N-P-A-N-N-A-A-I-V-L
M	Q-R-V-A-G-D-S-G-F-A-A-Y-S-R-Y
Patient 3	
Epitopes: (identified outliers)	
S2	S-P-D-V-D-L-G-D-I-S-G-I-N-A-S
N	T-A-S-W-F-T-A-L-T-Q-H-G-K-E-D

TABLE 3 Epitopes restricted to infected control patient

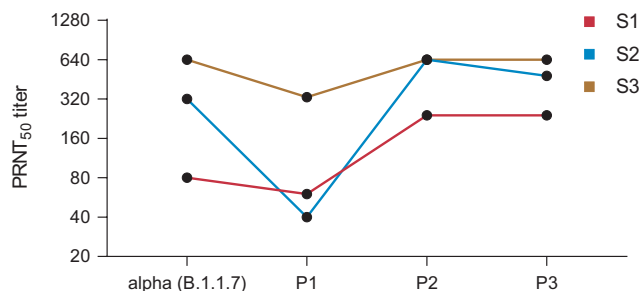| Control inf. patient 1		Control inf. patient 2		Control inf. patient 3	
Epitopes		**Epitopes**		**Epitopes**	
S1	H-K-N-N-K-S-W-M-E-S-E-F-R-V-Y	S1	C-T-F-E-Y-V-S-Q-P-F-L-M-D-L-E	S1	V-L-L-P-L-V-S-S-Q-C-V-N-L-T-T
	R-S-Y-L-T-P-G-D-S-S-S-G-W-T-A		R-S-Y-L-T-P-G-D-S-S-S-G-W-T-A		H-K-N-N-K-S-W-M-E-S-E-F-R-V-Y
	Q-P-R-T-F-L-L-K-Y-N-E-N-G-T-I	RBD	T-N-L-V-K-N-K-C-V-N-F-N-F-N-G		S-S-A-N-N-C-T-F-E-Y-V-S-Q-P-F
	T-D-A-V-D-C-A-L-D-P-L-S-E-T-K	S2	A-L-L-A-G-T-I-T-S-G-W-T-F-G-A		R-S-Y-L-T-P-G-D-S-S-S-G-W-T-A
RBD	G-P-K-K-S-T-N-L-V-K-N-K-C-V-N		L-S-R-L-D-K-V-E-A-E-V-Q-I-D-R	RBD	T-N-L-V-K-N-K-C-V-N-F-N-F-N-G
	F-N-F-N-G-L-T-G-T-G-V-L-T-E-S		L-Q-S-L-Q-T-Y-V-T-Q-Q-L-I-R-A		F-N-F-N-G-L-T-G-T-G-V-L-T-E-S
S1	S-V-I-T-P-G-T-N-T-S-N-Q-V-A-V	N	L-P-Y-G-A-N-K-D-G-I-I-W-V-A-T	S1	G-T-N-T-S-N-Q-V-A-V-L-Y-Q-D-V
	A-D-Q-L-T-P-T-W-R-V-Y-S-T-G-S				P-T-W-R-V-Y-S-T-G-S-N-V-F-Q-T
S2	L-S-R-L-D-K-V-E-A-E-V-Q-I-D-R			S2	V-T-L-A-D-A-G-F-I-K-Q-Y-G-D-C
	L-I-T-G-R-L-Q-S-L-Q-T-Y-V-T-Q				A-R-D-L-I-C-A-Q-K-F-N-G-L-T-V
	L-N-E-V-A-K-N-L-N-E-S-L-I-D-L				A-L-L-A-G-T-I-T-S-G-W-T-F-G-A
	Y-E-Q-Y-I-K-W-P-W-Y-I-W-L-G-F				L-S-S-N-F-G-A-I-S-S-V-L-N-D-I
M	T-V-E-E-L-K-K-L-L-E-Q-W-N-L-V				L-S-R-L-D-K-V-E-A-E-V-Q-I-D-R
	N-V-P-L-H-G-T-I-L-T-R-P-L-L-E				V-Q-I-D-R-L-I-T-G-R-L-Q-S-L-Q
					Y-E-Q-Y-I-K-W-P-W-Y-I-W-L-G-F
				N	L-G-T-G-P-E-A-G-L-P-Y-G-A-N-K
					L-P-Q-G-T-T-L-P-K-G-F-Y-A-E-G
				M	L-R-G-H-L-R-I-A-G-H-H-L-G-R-C

FIGURE 8 Plaque reduction neutralization test (PRNT) with sera of vaccinated subjects against isolates of long-term positive patients. Three sera from Comirnaty (BNT162b2, Biontech/Pfizer) vaccinated subjects (S1, S2, S3) were mixed in serial dilutions (1:20–1:640) with 75 PFU/ml of isolates P1, P2 and P3 for 1h at 37°C and inoculated to Vero E6 cells for 1 h at 37°C. Afterward infection plaques were visualized via crystal violet stain and counted. Alpha-variant B.1.1.7 was used as a control

The hypothesis of a potentially constrained efficiency of viral neutralization in all three patients is supported by characterization of linear epitopes recognized by antibodies of these patients (Figure 7; Table 2). No binding to linear epitopes within the RBD of the spike that is assumed to be crucial for neutralization was found.[44] Instead, linear epitopes derived from the nucleocapsid protein were recognized by sera from all three patients.

Plaque reduction neutralization tests revealed a potential neutralizing effect against infection with the three isolates similar to an infection with the alpha-variant (B.1.1.7) by sera of vaccinated subjects, indicating a protection by vaccination. Both the three isolates as well as the alpha-variant are characterized by a D614G mutation in the spike protein that was described as enhancing viral loads and improving transmission.[45,46] However, we observed differences in the efficiency of neutralization by sera of vaccinated individuals (Figure 8). The isolate P1 required less diluted sera to be blocked compared to the alpha-variant and the other isolates, although we detected no differences in gene sequences between P1 and P3. As the same PFU/ml was used in this approach, these results were unexpected.

Based on the data presented in this study, the assumption that genomic material detected by PCR in long-term positive patients is essentially remaining genome RNA fragments has to be critically reconsidered. Our findings provide strong evidence that long-term PCR-positive patients are possible carriers of still intact and infectious virus. Considering the long positivity of these patients, stopping quarantine periods after 10 days without PCR-based testing must be critically reconsidered because there might be a risk of virus transmission.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, E.H., T.J.M, S.V. and T.Z.; methodology, T.Z., I.M., S.H., J.R., C.M., Y.H., K.B. and D.B.; software, S.H., C.M.; formal analysis, T.Z., S.H. and C.M.; sample collection, T.Z., I.M., K.B. and D.B.; data curation, T.Z., T.J.M and E.H.; investigation, T.Z., I.M., S.H., J.R., C.M., Y.H., T.J.M., E.H.; writing—original draft preparation, T.Z., E.H., T.J.M., S.H., C.M.; writing—review and editing, T.Z., E.H. and T.J.M.; project administration, E.H., T.J.M., D.O., B.K-S. All authors have read and agreed to the published version of the manuscript.

INFORMED CONSENT STATEMENT

Informed consent was obtained from all subjects involved in the study.

ORCID

Sascha Hein https://orcid.org/0000-0001-7257-4538

REFERENCES

1. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506. 10.1016/S0140-6736(20)30183-5
2. Jiang S, Du L, Shi Z. An emerging coronavirus causing pneumonia outbreak in Wuhan, China: calling for developing therapeutic and prophylactic strategies. *Emerg Microbes Infect*. 2020;9(1):275-277. 10.1080/22221751.2020.1723441
3. Gorbalenya AE, Baker SC, Baric RS, et al. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. 2020; 5:536-544. https://doi.org/10.1038/s41564-020-0695-z
4. Wu A, Peng Y, Huang B, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe*. 2020;27(3):325-328. 10.1016/j.chom.2020.02.001
5. Zhou B, Yuan Y, Wang S, et al. Risk profiles of severe illness in children with COVID-19: a meta-analysis of individual patients. *Pediatr Res*. 2021;90(2):347-352. 10.1038/s41390-021-01429-2
6. Karagiannidis C, Mostert C, Hentschker C, et al. Case characteristics, resource use, and outcomes of 10 021 patients with COVID-19 admitted to 920 German hospitals: an observational study. *Lancet Respir Med*. 2020;8(9):853-862. 10.1016/S2213-2600(20)30316-7
7. Rabaan AA, Al-Ahmed SH, Al-Malkey M, et al. Airborne transmission of SARS-CoV-2 is the dominant route of transmission: droplets and aerosols. *Infez Med*. 2021;29(1):10-19.
8. Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*. 2020;581(7809):465-469. 10.1038/s41586-020-2196-x
9. Widders A, Broom A, Broom J. SARS-CoV-2: the viral shedding vs infectivity dilemma. *Infect Dis Health*. 2020;25(3):210-215. 10.1016/j.idh.2020.05.002
10. La Scola B, Le Bideau M, Andreani J, et al. Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards. *Eur J Clin Microbiol Infect Dis*. 2020;39(6):1059-1061. 10.1007/s10096-020-03913-9
11. Million M, Lagier J-C, Gautret P, et al. Early treatment of COVID-19 patients with hydroxychloroquine and azithromycin: a retrospective analysis of 1061 cases in Marseille, France. *Travel Med Infect Dis*. 2020;35:101738. 10.1016/j.tmaid.2020.101738
12. Wan X-F, Tang CY, Ritter D, et al. SARS-CoV-2 show no infectivity at later stages in a prolonged COVID-19 patient despite

positivity in RNA testing. *J Med Virol.* 2021;93(7):4570-4575. 10.1002/jmv.27001

13. Lu X, Wang L, Sakthivel SK, et al. US CDC real-time reverse transcription PCR panel for detection of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis.* 2020;26(8):1654-1665. 10.3201/eid2608.201246

14. Levin JZ, Yassour M, Adiconis X, et al. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat Methods.* 2010;7(9):709-715. 10.1038/nmeth.1491

15. Andrews S. FastQC: a quality control tool for high throughput sequence data. Bioinformatics babraham. 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

16. Zaharia M, Bolosky WJ, Curtis K et al. Faster and more accurate sequence alignment with SNAP. 2011. https://arxiv.org/pdf/1111.5572. Accessed July 10, 2021.

17. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-2120. 10.1093/bioinformatics/btu170

18. Morgan M, Pagès H, Obenchain V & Hayden N. Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import. R package version 2.8.0, 2021.

19. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079. 10.1093/bioinformatics/btp352

20. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics.* 2018;34(23):4121-4123. 10.1093/bioinformatics/bty407

21. Turoňová B, Sikora M, Schürmann C, et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science.* 2020;370(6513):203-208. 10.1126/science.abd5223

22. Amanat F, Stadlbauer D, Strohmeier S, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med.* 2020;26(7):1033-1036. 10.1038/s41591-020-0913-5

23. Díez-Fuertes F, Iglesias-Caballero M, García-Pérez J, et al. A founder effect led early SARS-CoV-2 transmission in Spain. *J Virol.* 2021;95(3):1-15. 10.1128/JVI.01583-20

24. Wang R, Chen J, Gao K, Hozumi Y, Yin C, Wei G-W. Analysis of SARS-CoV-2 mutations in the United States suggests presence of four substrains and novel variants. *Commun Biol.* 2021;4(1):228. 10.1038/s42003-021-01754-6

25. Ogawa J, Zhu W, Tonnu N, et al. The D614G mutation in the SARS-CoV2 spike protein increases infectivity in an ACE2 receptor dependent manner. *bioRxiv.* 2020. 10.1101/2020.07.21.214932

26. Rehman S, Mahmood T, Aziz E, Batool R. Identification of novel mutations in SARS-COV-2 isolates from Turkey. *Arch Virol.* 2020;165(12):2937-2944. 10.1007/s00705-020-04830-0

27. Sawatzki K, Hill N, Puryear W, Foss A, Stone J, & Runstadler J. *Ferrets not infected by SARS-CoV-2 in a high-exposure domestic setting.* 2021; Vol. 118(18):e2025601118. https://doi.org/10.1073/pnas.2025601118

28. Hodcroft EB, Zuber M, Nadeau S et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 202; *medRxiv.* 2021; 595(7869):707-712. 10.1038/s41586-021-03677-y

29. Yurkovetskiy L, Wang X, Pascal KE et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant; *Cell.* 2020;183(3):739-751.e8. 10.1016/j.cell.2020.09.032

30. Zuckerman NS, Bucris E, Drori Y, et al. Genomic variation and epidemiology of SARS-CoV-2 importation and early circulation in Israel. *PLoS One.* 2021;16(3):e0243265. 10.1371/journal.pone.0243265

31. Troyano-Hernáez P, Reinosa R, Holguín Á. Evolution of SARS-CoV-2 envelope, membrane, nucleocapsid, and spike structural proteins from the beginning of the pandemic to september 2020: a global and regional approach by epidemiological week. *Viruses.* 2021;13(2):243. 10.3390/v13020243

32. Li J-Y, Liao C-H, Wang Q, et al. The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. *Virus Res.* 2020;286:198074. 10.1016/j.virusres.2020.198074

33. Zinzula L. Lost in deletion: the enigmatic ORF8 protein of SARS-CoV-2. *Biochem Biophys Res Commun.* 2021;538:116-124. 10.1016/j.bbrc.2020.10.045

34. Zhang Y, Zhang J, Chen Y et al. The ORF8 protein of SARS-CoV-2 mediates immune evasion through potently downregulating MHC-I. *bioRxiv.* 2020; 10.1101/2020.05.24.111823

35. Hassan SS, Ghosh S, Attrish D et al. A unique view of SARS-CoV-2 through the lens of ORF8 protein. *Comput Biol Med.* 2020; 133:104380. 10.1016/j.compbiomed.2021.104380

36. Mateus J, Grifoni A, Tarke A, et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science.* 2020;370(6512):89-94. 10.1126/science.abd3871

37. Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J Med Virol.* 2020;92(5):495-500. 10.1002/jmv.25698

38. Noorimotlagh Z, Karami C, Mirzaee SA, Kaffashian M, Mami S, Azizi M. Immune and bioinformatics identification of T cell and B cell epitopes in the protein structure of SARS-CoV-2: a systematic review. *Int Immunopharmacol.* 2020;86:106738. 10.1016/j.intimp.2020.106738

39. Post N, Eddy D, Huntley C, et al. Antibody response to SARS-CoV-2 infection in humans: a systematic review. *PLoS One.* 2020;15(12):e0244126. 10.1371/journal.pone.0244126

40. Long Q-X, Tang X-J, Shi Q-L, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med.* 2020;26(8):1200-1204. 10.1038/s41591-020-0965-6

41. Li K, Huang B, Wu M, et al. Dynamic changes in anti-SARS-CoV-2 antibodies during SARS-CoV-2 infection and recovery from COVID-19. *Nat Commun.* 2020;11(1):6044. 10.1038/s41467-020-19943-y

42. Padoan A, Sciacovelli L, Basso D, et al. IgA-Ab response to spike glycoprotein of SARS-CoV-2 in patients with COVID-19: a longitudinal study. *Clin Chim Acta.* 2020;507:164-166. 10.1016/j.cca.2020.04.026

43. Sterlin D, Mathian A, Miyara M, et al. IgA dominates the early neutralizing antibody response to SARS-CoV-2. *Sci Transl Med.* 2021;13(577):1-10. 10.1126/scitranslmed.abd2223

44. Piccoli L, Park Y-J, Tortorici MA, et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell.* 2020;183(4):1024-1042.e21. 10.1016/j.cell.2020.09.037

45. Arora P, Pöhlmann S, Hoffmann M. Mutation D614G increases SARS-CoV-2 transmission. *Sig Transduct Target Ther.* 2021;6(1):101. 10.1038/s41392-021-00502-w

46. Plante JA, Liu Y, Liu J, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature.* 2021;592(7852):116-121. 10.1038/s41586-020-2895-3