A complete analysis of the epidemiological scenario around a SARS-CoV-2 reinfection: previous infection events and subsequent transmission

Laura Pérez Lago  
Gregorio Marañón General University Hospital  

Leire Pérez Latorre  
Gregorio Marañón General University Hospital  

Marta Herranz  
Gregorio Marañón General University Hospital  

Francisco Tejerina  
Gregorio Marañón General University Hospital  

Pedro J Sola Campoy  
Gregorio Marañón General University Hospital  

Jon Sicilia  
Gregorio Marañón General University Hospital  

Julia Suárez González  
Gregorio Marañón General University Hospital  

Cristina Andrés Zayas  
Gregorio Marañón General University Hospital  

Alvaro Chiner Oms  
Instituto de Biomedicina de Valencia-CSIC  

Santiago Jiménez Serrano  
Instituto de Biomedicina de Valencia-CSIC  

Neris García González  
Institute for Integrative Systems Biology (I2SysBio)  

Iñaki Comas  
Instituto de Biomedicina de Valencia-CSIC  

Fernando González Candelas  
Institute for Integrative Systems Biology (I2SysBio)  

Carolina Martínez Laperche  
Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM)  

Pilar Catalán  
Gregorio Marañón General University Hospital  

Patricia Muñoz
Research Article

Keywords: COVID-19, reinfection, transmission, epidemiology, genomics

DOI: https://doi.org/10.21203/rs.3.rs-106167/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The first descriptions of reinfection by SARS-CoV-2 have been recently reported. However, these studies focus exclusively on the reinfected case, without considering the epidemiological context of the event. We present the first complete analysis of the epidemiological scenario around a reinfection by SARS-CoV-2, including three cases preceding the reinfection, the reinfected case per se, and the subsequent transmission to another seven cases. Our analysis is supported by host genetics, viral whole genome sequencing, phylogenomic population analysis, and refined epidemiological data obtained from in-depth interviews with the involved subjects. The reinfection involved a 53-year-old woman with asthma, with a first COVID-19 episode in April 2020 and a much more severe second episode four months and a half later, with COVID-19 seroconversion in August, and requiring hospital admission.

Introduction

The first case of a SARS-CoV-2 reinfection was reported on August 24, 2020, in a patient from Hong Kong who acquired a second infection after having travelled to Europe, where he was exposed to a strain-lineage different from the one involved in the first episode.

Since this first report, few other COVID-19 reinfections have been published or deposited in repositories. The age of the affected subjects, the time between episodes, and the severity of the second episode with respect to the first one are rather variable, which makes it difficult to determine a common reinfection pattern.

To distinguish between prolonged SARS-CoV-2 shedding -found to extend up to 101 days in some cases and reinfection, requires whole genome sequencing (WGS) strategies to determine differences between the first and second strains. Changes between strains may be notable, e.g., when they correspond to different lineages, or moderate, yet enough to demonstrate that each strain followed a different evolutionary path. Moreover, the European Centre for Disease Prevention and Control (ECDC) has recently suggested the use of WGS to also document reinfections by demonstrating that the strain involved in the reinfection is clustered with other strains circulating in the location of the exposure.

To date, published reinfection cases only focus on the analysis of the case per se. Here, we describe a new SARS-CoV-2 reinfection case without clinical risk factors and analyse the epidemiological scenario before and after the reinfection. We identify the exposure event responsible for the reinfection and describe onward extensive transmission started from the reinfected case.

Materials And Methods

Diagnostic RT-PCR

RNA was extracted and purified from 300μL of nasopharyngeal exudates by using the EasyMag (Biomeriux, France)(specimen from April) or KingFisher (Thermo Fisher Scientific, Waltham,
Massachusetts) (specimens from August/September) equipments.

Next, an RT-PCR was performed, using in the April specimen the Novel Coronavirus (2019-nCoV) Nucleic Acid Diagnostic Kit (Sansure Biotech, China) and in the August/September specimens the TaqPath COVID-19 CE-IVD RT-PCR kit (Thermo Fisher Scientific, USA).

**Whole genome sequencing**

11 μL of RNA were used as template for reverse transcription using Invitrogen SuperScript IV reverse transcriptase (ThermoFisher Scientific, Massachusetts, USA) and random hexamers (ThermoFisher Scientific, Massachusetts, USA). Whole genome amplification of the coronavirus was done using the Artic_nCov-2019_V3 panel of primers (Integrated DNA Technologies, Inc., Coralville, Iowa, USA) (artic.network/ncov-2019) and the Q5 Hot Start DNA Polymerase enzyme (New England Biolabs, Ipswich, Massachusetts, USA). Libraries were prepared using Nextera Flex DNA Library Preparation Kit (Illumina Inc, California, USA) following the manufacturer's instructions.

Libraries were quantified using the Quantus™ Fluorometer (Promega, Wisconsin, USA), before being pooled at equimolar concentration (4 nM). Next, libraries were sequenced in pools of up 17 libraries on the Miseq system (Illumina Inc, California, USA) using the MiSeq Reagent Micro kit v2 (2x151pb) or in pools of up to 96 libraries with the MiSeq Reagent (2x201 pb).

FastQ files above the GISAID thresholds were deposited at GISAID (hCoV-19/Spain/MD-IBV-99007733/2020, hCoV-19/Spain/MD-IBV-99007151/2020, hCoV-19/Spain/MD-IBV-99007734/2020 and hCoV-19/Spain/MD-IBV-99007170/2020). All sequences were also deposited at ENA (European Nucleotide Archive; https://www.ebi.ac.uk/ena/browser/home)(ERS5219974, ERS5220028, ERS5219973, ERS5219975, ERS5220046).

A bioinformatics pipeline developed by the SeqCOVID consortium was applied was applied to analyse the sequencing reads. The pipeline is based on the computational tool iVAR [https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1618-7], which can be accessed at https://gitlab.com/sabio-NGS/sars-cov2-mapping. Briefly, the pipeline goes through the following steps: 1) removal of human reads with Kraken [https://genomebiology.biomedcentral.com/articles/10.1186/gb-2014-15-3-r46]; 2) pre-processing of fastq files using fastp [https://academic.oup.com/bioinformatics/article/34/17/i884/5093234] v 0.20.1 (arguments: --cut tail, --cut-window-size, --cut-mean-quality , -max_len1 ,-max_len2 ); 3) mapping and variant calling using IVAR version 1.2; 4) quality control assessment with MultiQC[http://dx.doi.org/10.1093/bioinformatics/btw354].

The consensus sequences obtained with the application of this pipeline were aligned against the SARS-CoV-2 reference sequence [https://pubmed.ncbi.nlm.nih.gov/32015508/] using MAFFT [https://academic.oup.com/mbe/article/30/4/772/1073398]. Problematic positions were masked using the mask_alignment.py script from the repository maintained by Rob Lanfear.
The clade was assigned according to the Nextstrain nomenclature. A maximum-likelihood phylogenetic tree was reconstructed including all the masked positions from the sequences along with sequences obtained similarly from the same population in Madrid corresponding to the health district from the case, uploaded to GISAID up to October 8, 2020. We used IQtree v2, with model GTRHKY+G as substitution model and czb option, and rooted the tree to with the a basal strain from Wuhan-1 reference sequence.

**Short tandem repeat analysis**

Human identity testing analysis was performed by short tandem repeat (STR)-PCR (Mentype® Chimera® Biotype, Germany) on the same specimens used to perform the SARS-CoV-2 RT-PCR and which were also sequenced. We examined 12 non-coding STR loci and the gender-specific locus amelogenin (Supplementary Table), labelled with three different dyes (6-FAM™, BTG, or BTY). The selected loci offer a very high rate of heterozygosity and a balanced allelic distribution (Thiede et al., 2004). PCR was performed with 0.2-1 ng of genomic DNA using the Mentype® Chimera® PCR amplification kit (Biotype, Germany), the GeneAmp® PCR System 9700 Thermal Cycler (Applied Biosystems), and subsequent capillary electrophoresis in a Genetic Analyzer 3130xl (Applied Biosystems) under conditions recommended by the manufacturer.

**Results**

A 53-year-old woman (Case A) was admitted to the emergency room in our hospital (Gregorio Marañón, Madrid, Spain) on April 3, 2020 due to dyspnoea, fever, cough with expectoration of 24 hours of evolution, and history of bronchial asthma. Blood tests and chest X-ray were ordered and results showed no outstanding changes. The SARS-CoV2 PCR on nasopharyngeal exudate was positive (Ct value 30). Since the patient had no respiratory failure nor other serious conditions, she was discharged from hospital the same day and given symptomatic treatment. The patient remained symptomatic for over two weeks (mainly dyspnoea and fever) with eventual resolution of symptoms. On April 18, a second RT-PCR was performed with a positive result. One month after discharge, the SARS-CoV 2 RT-PCR was repeated (May 12) and this time the result was negative. COVID-19 serology was not available at the time.

The source of infection was undetermined. Case A reported having been confined to her home with her husband during the three weeks prior to the beginning of symptoms. She denied contact with anyone else during confinement. Her husband did not have any symptoms and therefore no RT-PCR was performed.

On August 14, four months and a half (140 days) from her first positive RT-PCR, Case A began to have fever, dyspnoea, cough, and arthromyalgia. Another SARS-CoV2 RT-PCR was performed (August 21) with a positive result (Ct value 22).
On August 25, Case A went back to the emergency room, this time with respiratory failure and multiple bilateral pulmonary infiltrates and was admitted to the hospital. No SARS-Cov-2 antibodies were detected in the admission tests. Upon admission, the presence of lymphopenia, mild hypertransaminasemia, and elevated LDH and CRP stood out.

During the first 48 hours, Case A showed radiological worsening and respiratory failure, exhibiting significant bronchospasm. She was given corticosteroids, remdesivir, and lopinavir/ritonavir. Three RT-PCRs were performed during hospitalization (August 28, and September 2 and 6, all positive (Ct values 21, 33, and 33 respectively). The patient progressively improved, and 17 days after admission was discharged with oxygen therapy. Prior discharge, the COVID-19 serology (SARS-CoV-2 IgG Architect, Abbott, Chicago USA) was repeated (8/9). The result was positive with titers of 7.04.

**Epidemiological events preceding reinfection**

Twelve days before (August 7) Case A´s second episode, she had had close contact with an uncle (without wearing facemasks, tight physical contact, including kisses and hugs, Figure 1). Besides Case A, no other member in the family participated in that contact with the uncle. Two days later, (August 9) the uncle developed a cough, arthromyalgias, asthenia, and dysthermia. The RT-PCR-SARS-CoV-2 test performed on August 12 was positive (Ct value 19). His symptoms resolved on August 18.

Case A was interviewed to obtain more details from her and her uncle´s epidemiological context. Three of her uncle´s friends had also fallen ill with COVID-19; all four attended the same mosque every Friday and had had additional contacts. One of them was interviewed and referred that the last time all four coincided was at the mosque for the Friday ceremony on August 7. The three friends had subsequent positive RT-PCRs (August 11, 15 and 16; Figure 1). Case A, her husband and daughter had not attended that mosque or met the uncle´s friends.

**Epidemiological events following reinfection**

On August 25, eleven days after the onset of Patient Case A´s symptoms in the second episode, her husband started to report fever and malaise. He tested positive for SARS-CoV2 PCR (Ct value 25). On August 28 his dyspnoea worsened and an interstitial infiltrate was observed in the upper left lobe. Home isolation was indicated for 14 days.

On August 18 Case A´s daughter, who visited her daily, started with a cough odynophagia, asthenia, and fever of 39 º C. On August 21, she had a positive RT-PCR. Home isolation ended after 14 days, without an RT-PCR control. The daughter´s husband began having symptoms the same day (August 18) and had a positive RT-PCR on August 21 and a second one on September 2. The RT-PCR tests performed on their four children were all positive, but only one of them developed symptoms (starting on August 22) (Figure 1).

**Genomic analysis**
We first confirmed that the specimens isolated from Case A, who had had positive SARS-CoV-2 RT-PCR results in April (first episode) and August (second episode) belonged to the same patient, as indicated by the identical microsatellite STR-PCR patterns obtained from the human DNA in the corresponding samples (Supplementary Figure 1).

Next, we performed comparative analyses of the SARS-CoV-2 sequences of the strains isolated from Case A in the first and second episodes. The analysis of Episode 1 specimen allowed us to confirm that sequences corresponded to SARS-CoV-2, but it did not offer enough coverage (only 18% of the chromosome offered at least 30X coverage) to determine the complete consensus sequence and call SNPs with high confidence. WGS analysis of Episode 2 specimen provided good coverage (99% of the genome with >30X coverage depth and a total of 790545 mapped reads) and allowed us to determine 16 SNPs relative to the Wuhan-1 reference (7 of them missense variants) (Figure 2) and assign the lineage of the strain (20A).

Next, we decided to extend the genomic analysis beyond Case A, to include i) the strain from the potential source of Case A’s second episode, her uncle, and ii) the strains from two of the uncle’s friends, representing his exposure context. In addition, we included Case A’s husband strain as potential subsequent receptor from Case A. No samples were available for sequencing from Case A’s daughter, her son-in-law, or her granddaughter. The four specimens yielded sequences of enough quality (>93-99% of the genome with >30X coverage depth and 152354-525291 mapped reads; and in one specimen 77.53% and 290438) to allow comparisons throughout the complete genome. Case A and her husband presented identical sequences and both differed from those of case A’s uncle and his friends in 1 SNP (Figure 2). These data indicate that the strain involved in Case A’s reinfection was circulating in the epidemiological context of Case A’s uncle. The acquisition of one SNP, the presence of this SNP shared by Case A and her husband, the chronology of the cases, and the fact that Case A husband did not attend the mosque or met Case A’s uncle nor his friend, suggest a direction of transmission from the uncle to Case A and then from Case A to her husband.

An extended phylogenetic analysis, including 348 strains sequenced in our institution from the same population in Madrid, showed that this strain was circulating by the time Case A suffered the second episode, August/September 2020. The strain was part of a clade including exclusively strains after June 2020 (blue clade) and the five cases in this study shared a single proper branch within this clade (Figure 3). The strain, or related strains, were not found among the strains circulating in the same population in Madrid at the time range corresponding to Case A’s first episode (end of March/beginning of April; red clades; Figure 3). In agreement, a dating analysis available in nextstrain.org shows that the second episode strain belongs to a clade, 20A.EU1, that had no representative sample before the end of June 2020 (Supplementary Figure 2).

**Discussion**
Reinfection as the most likely explanation for the second episode of our patient (Case A) is based on: i) four months and a half between her first and second episodes and confirmation by STR analysis that the specimens from the two positive SARS-CoV-2 RT-PCR episodes came from the same patient, ii) negative PCR after the resolution of the first episode, iii) COVID-19 antibody seroconversion in Case A in the second episode, iv) close physical contact without protection with a family member (uncle) who began with symptoms two days after their encounter, had a positive RT-PCR five days later, and started with symptoms five days before the onset of Case A’s symptoms in her second episode, v) identification by WGS that SARS-CoV-2 strains from Case A and her uncle’s were almost identical (1 SNP), and vi) determination that the strain was circulating in her uncle’s epidemiological close context and in Madrid during the second episode, but not at the time of the first episode.

Case A did had no evidence of immunosuppression. She only received regular treatment with inhaled budesonide because of her asthma with no systemic corticosteroids between April and August 2020. At admission, HIV serology and a study of IgG, IgM and IgA were requested; negative and normal results were obtained, respectively. A similar reinfection case was reported in Belgium; the patient, was a women, immunocompetent, similar age (51 years), and had asthma as the only risk factor, for which she was given oral corticosteroids.

Time lapses between the first and second episodes in previously reported COVID-19 reinfection cases vary considerably, from 19 days (the shorter) to 142 (the longest). The case reported in this study constitutes the second longest period between the first and second episode, 140 days.

To date there is no defined severity pattern for second episodes following reinfection. Some of these second episode reinfection cases have been reported to be asymptomatic, while others presented with milder or more severe symptoms. In the case described in this study, the second episode was much more severe, causing pneumonia and requiring hospital admission.

In most reinfection cases, reported COVID-19 IgG serology was not well documented for the first episodes because at that time (March/April 2020) it was not routinely performed. This implies that in some reinfections with seroconversion in the second episode, as in our case, it is not possible to determine whether the case did not mount immune response in the first episode or antibodies were lost before the second episode was diagnosed. Similar uncertainties have been described for SARS-CoV-2 reinfections elsewhere.

Genomics is the ultimate support to document reinfections. The most obvious way to determine a reinfection in studies already published or reported is to compare the strains from each episode. Unfortunately, and similarly to other COVID-19 reinfection reports, we could not obtain enough sequence information from the sample analyzed in the first episode, to conclusively show that it was different from that from the second episode.
However, we further confirmed by genomic analysis and epidemiological enquiry that Case A`s second episode was caused by a strain that came from her uncle, the case involved in the high-risk exposure causing her reinfection. To the best of our knowledge, this is the first study that identifies the strain involved in a reinfection (Case A in this study) among subjects in close context, preceding the case. Recently, ECDC has accepted that reinfections may be documented either by identifying genomic differences between the two episodes or by confirming that the strain from the second episode clusters with strains from the location of exposure. In this study, we not only identified the strain responsible for the reinfection of Case A (isolated from her uncle) but also proved its circulation in his epidemiological context in three other subjects, preceding the reinfection of Case A. Furthermore, we performed an extended phylogenetic analysis and determined that the strain was part of a clade of strains circulating in the same population in Madrid during the period Case A suffered a second episode, but not during her first episode. In fact, this clade had no representative sample, at a global collection, before June 2020 and it corresponded to the clade 20A-EU1, which after emerging in Spain at the end of June spread successfully through Europe. The reinfection event, involving five cases, shared an independent branch within the clade.

The carrier (uncle) of the strain causing Case A`s reinfection was asymptomatic at the time they had close contact; his symptoms began two days later. Several studies have demonstrated that SARS-CoV-2 can be transmitted before the onset of symptoms.

Understanding Case A`s epidemiological context allowed us to show for the first time the onward transmission leading to a reinfection and to report subsequent transmission from a reinfection case. The identity of Case A and her husband`s strains, including a private marker SNP not shared by other cases, the chronology of symptom onset, and the confirmation that Case A`s husband did not interacted with her uncle nor his friends, and did not attend the mosque, confirmed Case A as the sole source for the infection of her husband.

The magnitude of the transmission from reinfevted Case A, beyond her husband, was notorious, as can be deduced from the sequential chronology of symptoms and positive RT-PCRs obtained for her daughter, son-in-law, and four grandsons. Unfortunately, no samples were was available from these subjects as they were diagnosed in another institution where samples were not stored.

Here we describe the first complete analysis of the epidemiological scenario around a reinfection by SARS-CoV2 supported on host genetic analysis, viral genomic analysis, phylogenomic population analysis, and in-depth epidemiological investigation. This extensive approach documents a much more severe second COVID-19 episode in a woman without risk factors, except for asthma, after exposure through a close contact to a SARS-CoV-2 strain actively transmitted in the same setting/population. Once reinfevted, the viable virus from this case caused an extensive subsequent transmission among family members.

**Declarations**
Acknowledgements

This work was supported by Instituto de Salud Carlos III (Ref COV20/00140: SeqCOVID - Consorcio para la epidemiología genómica de SARS-CoV-2 en España) and by Consejo Superior de Investigaciones Científicas (CSIC) (PTI Salud Global) . Miguel Servet Contract CP15/00075) to LPL. We are grateful to Dainora Jaloveckas (cienciaturaducida.com) for editing and proofreading assistance.

There is no conflict of interests for any of the authors.

Transparency declaration

The lead author (Darío García de Viedma) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Ethical guidelines

All appropriate ethical guidelines for the use of human subjects as required by the Ethics Committee in our Institution have been followed. Patient consents, including the authorization by the patients for publishing the case have been obtained and archived.

Gregorio Marañón Microbiology-ID COVID 19 Study Group

Adán-Jiménez (Javier), Alcalá (Luis), Aldámiz (Teresa), Alonso (Roberto), Álvarez (Beatriz), Álvarez-Uría (Ana), Arias (Alexi), Arroyo (Luis Antonio), Berenguer (Juan), Bermúdez (Elena), Bouza (Emilio), Burillo (Almudena), Candela (Ana), Carrillo (Raquel), Catalán (Pilar), Cercenado (Emilia), Cobos (Alejandro), Díez (Cristina), Escribano (Pilar), Estévez (Agustín), Fanciulli (Chiara), Galar (Alicia), García (Mª Dolores), García de Viedma (Darío), Gijón (Paloma), González (Adolfo), Guillén (Helmuth) Guineo (Jesús), Haces (Laura Vanessa), Herranz (Marta), Kestler (Martha), López (Juan Carlos), Losada (Carmen Narcisa), Machado (Marina), Marín (Mercedes), Martín (Pablo), Montilla (Pedro), Moure (Zaira), Muñoz (Patricia), Olmedo (María), Padilla (Belén), Palomo (María), Parras (Francisco), Pérez-Granda (María Jesús), Pérez-Lago (Laura), Pérez (Leire), Pescador (Paula), R Maus (Sandra), Reigadas (Elena), Rincón (Cristina), Rodríguez (Belén), Rodríguez (Sara), Rodríguez-Grande (Cristina), Rojas (Adriana), Ruiz-Serrano (María Jesús), Sánchez (Carlos), Sánchez (Mar), Serrano (Julia), Sola Campoy (Pedro J), Tejerina (Francisco), Valerio (Maricela), Veintimilla (Mª Cristina), Vesperinas (Lara), Vicente (Teresa), de la Villa (Sofía).

References


